

09/890, 806.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 11:37:25 ; Search time 21.7747 Seconds
(without alignments)
225.401 Million cell updates/sec

Title: US-09-890-806-3_COPY_28_143
Perfect score: 629
Sequence: 1 AGEDALRPWKSTAKHPWFOI.....LQGDVIMRCVPELRVDYTS 116

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	77.5	12.3	676	3	US-08-947-965-71 Sequence 71, Appl
2	73.5	11.7	729	4	US-09-252-991A-32535 Sequence 32535, A
3	70.5	11.2	675	3	US-08-947-965-76 Sequence 76, Appl
4	69	11.0	993	4	US-09-514-302-4 Sequence 4, Appli
5	69	11.0	1938	4	US-09-514-302-2 Sequence 2, Appli
6	68.5	10.9	640	4	US-09-252-991A-27542 Sequence 27542, A
7	67	10.7	240	4	US-09-152-060-100 Sequence 100, App
8	67	10.7	335	4	US-09-152-060-64 Sequence 64, Appl
9	67	10.7	391	3	US-08-706-216-6 Sequence 6, Appli
10	66.5	10.6	100	4	US-09-430-702-1 Sequence 1, Appli
11	66	10.5	790	3	US-08-960-780-4 Sequence 4, Appli
12	66	10.5	790	3	US-09-073-898-4 Sequence 4, Appli
13	65.5	10.4	516	4	US-08-868-373-14 Sequence 14, Appl
14	65	10.3	587	4	US-09-328-352-8096 Sequence 8096, Ap
15	64.5	10.3	310	4	US-09-328-352-8096 Sequence 17381, A
16	64	10.2	746	2	US-08-838-219B-6 Sequence 6, Appli
17	64	10.2	746	3	US-09-233-336A-6 Sequence 6, Appli
18	64	10.2	746	3	US-09-233-336A-6 Sequence 6, Appli
19	64	10.2	746	3	US-09-402-036-6 Sequence 6, Appli
20	64	10.2	746	4	US-09-904-226-6 Sequence 6, Appli
21	64	10.2	789	1	US-08-471-033-29 Sequence 29, Appl
22	64	10.2	789	1	US-08-471-033-32 Sequence 32, Appl
23	64	10.2	789	2	US-08-471-044-29 Sequence 29, Appl
24	64	10.2	789	2	US-08-471-044-32 Sequence 32, Appl
25	64	10.2	789	2	US-08-463-483A-29 Sequence 29, Appl
26	64	10.2	789	2	US-08-463-483A-32 Sequence 32, Appl
27	64	10.2	789	2	US-08-471-046A-29 Sequence 29, Appl

28	64	10.2	789	2	US-08-471-046A-32	Sequence 32, Appl
29	64	10.2	789	2	US-08-470-566B-29	Sequence 29, Appl
30	64	10.2	789	2	US-08-470-566B-32	Sequence 32, Appl
31	64	10.2	789	2	US-08-838-219B-2	Sequence 2, Appli
32	64	10.2	789	2	US-08-838-219B-4	Sequence 4, Appli
33	64	10.2	789	2	US-08-469-334-29	Sequence 29, Appl
34	64	10.2	789	2	US-08-469-334-32	Sequence 32, Appl
35	64	10.2	789	3	US-09-300-529-29	Sequence 29, Appl
36	64	10.2	789	3	US-09-300-529-32	Sequence 32, Appl
37	64	10.2	789	3	US-09-233-336A-4	Sequence 2, Appli
38	64	10.2	789	3	US-09-233-336A-2	Sequence 2, Appli
39	64	10.2	789	3	US-09-233-752A-2	Sequence 2, Appli
40	64	10.2	789	3	US-09-233-752A-4	Sequence 4, Appli
41	64	10.2	789	3	US-09-402-036-2	Sequence 2, Appli
42	64	10.2	789	3	US-09-402-036-4	Sequence 4, Appli
43	64	10.2	789	4	US-09-002-285-78	Sequence 78, Appl
44	64	10.2	789	4	US-09-002-285-80	Sequence 80, Appl
45	64	10.2	789	4	US-09-002-285-92	Sequence 92, Appl

ALIGNMENTS

RESULT 1
US-08-947-965-71
; Sequence 71, Application US/08947965A
; Patent No. 6004790
; GENERAL INFORMATION:
; APPLICANT: Dijkhuizen, Lubbert
; APPLICANT: Dijkstra, Bauke
; APPLICANT: Andersen, Carsten
; APPLICANT: Oster, Claus von der
; TITLE OF INVENTION: Cyclomaltohextrin Glucanotransferase
; TITLE OF INVENTION: Variants
; FILE REFERENCE: 4285.204-US
; CURRENT APPLICATION NUMBER: US/08/947, 965A
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 0477/95
; EARLIER FILING DATE: 1995-04-21
; EARLIER APPLICATION NUMBER: 1173/95
; EARLIER FILING DATE: 1995-10-17
; EARLIER APPLICATION NUMBER: 1281/95
; EARLIER FILING DATE: 1995-11-16
; EARLIER APPLICATION NUMBER: PCT/DK96/00179
; EARLIER FILING DATE: 1996-04-22
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 71
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-08-947-965-71

Query Match 12.3%; Score 77.5; DB 3; Length 676;
Best Local Similarity 24.0%; Pred. No. 0.67;
Matches 25; Conservative 22; Mismatches 44; Indels 13; Gaps 3;
QY 12 TAKHPWFQIEDNRKCYIDNGKLFARGSIGVMSRFVDPKADYGGVGENLYVHADVEFVP 71
DB 132 TPNHSSPALETPNVYENGAIYNGALLGN-----YSNDQQLFHHNGTDFSS 180
QY 72 GE-SLKMNVRNLDVMPFETLALRLVLCQDVIWL-RCVPELRVD 113
DB 181 YEDSIYRNLYDLADYDLNNTVMQYLKESIKFWLDGIDGIRD 224
RESULT 2
US-09-252-991A-32535
; Sequence 32535, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 32535
;; LENGTH: 729
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32535

Query Match 11.7%; Score 73.5; DB 4; Length 729;
Best Local Similarity 29.3%; Pred. No. 2.3;
Matches 34; Conservative 16; Mismatches 41; Indels 25; Gaps 7;

QY 1 AGED-----ALRPWKSTAKHPFQIE-DNRCYIDNGKLFARGSIVGNMSRFVDPKADYG 54
DB 559 AGEDLGVAALVDPARTHVROALAQVDLRLVGI-----RAGGVDEHRRVFFAAEGG-G 611
QY 55 GYGENLYVHAD-DYEFVPGESLKMNVNLDVMPIFE---TLALRLVLQGDVYIWLRC 106
DB 612 GYGEEDFFHRDADVRARAGQ-----VDLARVERLRRLAVELLKAGDVYLLRC 659

RESULT 3
US-08-947-965-76
; Sequence 76, Application US/08947965A
; Patent No. 6004790
; GENERAL INFORMATION:
; APPLICANT: Dijkhuizen, Lubbert
; APPLICANT: Dijkstra, Bauke
; APPLICANT: Andersen, Carsten
; APPLICANT: Osten, Claus von der
; TITLE OF INVENTION: Cyclomaltoedextrin Glucanotransferase
; TITLE OF INVENTION: Variants
; FILE REFERENCE: 4285.204-US
; CURRENT APPLICATION NUMBER: US/08/947,965A
; CURRENT FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 0477/95
; EARLIER FILING DATE: 1995-04-21
; EARLIER APPLICATION NUMBER: 1173/95
; EARLIER FILING DATE: 1995-10-17
; EARLIER APPLICATION NUMBER: 1281/95
; EARLIER FILING DATE: 1995-11-16
; EARLIER APPLICATION NUMBER: PCT/DK96/00179
; EARLIER FILING DATE: 1996-04-22
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastsEQ for windows Version 3.0
; SEQ ID NO 76
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Bacillus ohbensis
US-08-947-965-76

Query Match 11.2%; Score 70.5; DB 3; Length 675;
Best Local Similarity 25.0%; Pred. No. 4.8;
Matches 26; Conservative 21; Mismatches 44; Indels 13; Gaps 4;

QY 12 TAKHPWFQIEDNRCYIDNGKLFARGSIVGNMSRFVDPKADYGVGENLYVHADVEFPV 71
DB 130 TPNHSSPALETPDPSYAENGAVVNDGVLIGNYSN---DP-----NNLFHHNGGTDFSS 178
QY 72 GE-SLKMNVNLDVMPIFEFTLALRLVLQGDVYIWL-RCVPELRVD 113
DB 179 YEDSIYRNLYDLADYDLNNTVWMDQYLKESIKLWLDKGIDGIRVD 222

RESULT 4
US-09-514-302-4

;; Sequence 4, Application US/09514302
;; Patent No. 6338959
;; GENERAL INFORMATION:
;; APPLICANT: HATADA, Yuji
;; APPLICANT: IGARASHI, Kazuaki
;; APPLICANT: OZAKI, Katsuya
;; APPLICANT: ARA, Katsutoshi
;; APPLICANT: KAWAI, Shuji
;; APPLICANT: ITO, Susumu
;; TITLE OF INVENTION: GENE FOR ENZYME HAVING BOTH ALKALINE PULLULANASE AND
;; TITLE OF INVENTION: ALKALINE ALPHA-AMYLASE ACTIVITIES
;; FILE REFERENCE: 2173-105P
;; CURRENT APPLICATION NUMBER: US/09/514,302
;; CURRENT FILING DATE: 2000-02-28
;; EARLIER APPLICATION NUMBER: 08/952,084
;; EARLIER FILING DATE: 1997-11-10
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 893
;; TYPE: PRT
;; ORGANISM: Bacillus sp.
US-09-514-302-4

Query Match 11.0%; Score 69; DB 4; Length 893;
Best Local Similarity 32.0%; Pred. No. 11;
Matches 24; Conservative 11; Mismatches 30; Indels 10; Gaps 3;

QY 12 TAKHPWFQIEDNRCYIDNGKLFARGSIVGNMSRFVDPKAD-----YGGVGENLYVHAD 65
DB 82 SVKSGWKLIDEMYAY--DGKLGAEIHEHDGTATLKWSPKADNVSVLYDKVDQNEVV--D 137
QY 66 DVEFVPGESLKMNV 80
DB 138 TIEMVKGDGRGWSVK 152

RESULT 5
US-09-514-302-2
; Sequence 2, Application US/09514302
; Patent No. 6338959
; GENERAL INFORMATION:
; APPLICANT: HATADA, Yuji
; APPLICANT: IGARASHI, Kazuaki
; APPLICANT: OZAKI, Katsuya
; APPLICANT: ARA, Katsutoshi
; APPLICANT: KAWAI, Shuji
; APPLICANT: ITO, Susumu
; TITLE OF INVENTION: GENE FOR ENZYME HAVING BOTH ALKALINE PULLULANASE AND
; TITLE OF INVENTION: ALKALINE ALPHA-AMYLASE ACTIVITIES
; FILE REFERENCE: 2173-105P
; CURRENT APPLICATION NUMBER: US/09/514,302
; CURRENT FILING DATE: 2000-02-28
; EARLIER APPLICATION NUMBER: 08/952,084
; EARLIER FILING DATE: 1997-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1938
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-514-302-2

Query Match 11.0%; Score 69; DB 4; Length 1938;
Best Local Similarity 32.0%; Pred. No. 31;
Matches 24; Conservative 11; Mismatches 30; Indels 10; Gaps 3;

QY 12 TAKHPWFQIEDNRCYIDNGKLFARGSIVGNMSRFVDPKAD-----YGGVGENLYVHAD 65
DB 1127 SVKSGWKLIDEMYAY--DGKLGAEIHEHDGTATLKWSPKADNVSVLYDKVDQNEVV--D 1182
QY 66 DVEFVPGESLKMNV 80

RESULT 9

US-08-706-216-6

; Sequence 6, Application US/08706216
; Patent No. 6140098

; GENERAL INFORMATION:

APPLICANT: Balasubramanian, Sriam
APPLICANT: Ford, John
APPLICANT: Gorman, Daniel M.
APPLICANT: Zurawski, Gerard
TITLE OF INVENTION: MAMMALIAN PROTEASES; RELATED REAGENTS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNA Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,216
FILING DATE: 30-AUG-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-9196
TELEFAX: 415-496-1200

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 391 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-706-216-6

Query Match 10.7%; Score 67; DB 3; Length 391;

Best Local Similarity 26.4%; Pred. No. 6.2;

Matches 29; Conservative 18; Mismatches 35; Indels 28; Gaps 6;

QY 14 KHPWFQIED-----NRCYIDNCKLFARSGIVGNMRFVDPKADYGVGEN----- 59

DB 151 RHAVYQAEHLQTAGCGVSDSL---GSLGPRTA VFRP-----GDSLPSRETRYV 202

QY 60 -LYVHADVVEF-VPGESLKMVNVNLDVM---PIFETLALRLVLCGDVIV 103

DB 203 ELYVVVDNAEFQMLGSEAAVRRHVLEVVNHVDKLYQKLNFRVVLVGLHEIW 252

RESULT 10

US-09-430-702-1

; Sequence 1, Application US/09430702

; Patent No. 6472192

; GENERAL INFORMATION:

APPLICANT: SCHULZ, GEORG E.
APPLICANT: PARSIEGLA, GOETZ
APPLICANT: CANDUSSIO, ANTON
APPLICANT: WICH, GUNTER

TITLE OF INVENTION: CYCLODEXTRIN GLYCOSYL TRANSFERASES FOR PRODUCING
TITLE OF INVENTION: Y-CYCLODEXTRIN
FILE REFERENCE: SCHULZ-W2 CIP

CURRENT APPLICATION NUMBER: US/09/430,702

EARLIER APPLICATION NUMBER: 08/816,317

EARLIER FILING DATE: 1997-03-13

NUMBER OF SEQ ID NOS: 44

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 1

; LENGTH: 100

; TYPE: PRT

; ORGANISM: Bacillus sp. 1-1

US-09-430-702-1

Query Match 10.6%; Score 66.5; DB 4; Length 100;

Best Local Similarity 25.9%; Pred. No. 1.1;

Matches 15; Conservative 10; Mismatches 22; Indels 11; Gaps 1;

QY 12 TAKHPWFQIEDNRCYIDNCKLFARSGIVGNMRFVDPKADYGVGENLYVHADVVEF 69

DB 40 TPNHSPALETPNVYENCAIYDNCALLGN-----YENDQNLPHHNGTIN 86

RESULT 11

US-08-960-780-4

; Sequence 4, Application US/08960780

; Patent No. 6204435

; GENERAL INFORMATION:

APPLICANT: Feltelson, Jerald S.
APPLICANT: Schneck, H. Ernest
APPLICANT: Narva, Kenneth E.
APPLICANT: Stockhoff, Brian A.
APPLICANT: Schmeits, James

APPLICANT: Loewer, David

APPLICANT: Dullum, Charles Joseph

APPLICANT: Muller-Cohn, Judy

APPLICANT: Stamp, Lisa

TITLE OF INVENTION: No. 6204435e1 Pesticidal Toxins and Nucleotide

NUMBER OF SEQUENCES: 134

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: US

ZIP: 32606-6669

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/960,780

FILING DATE: 30-OCT-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/029,848

FILING DATE: 30-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: MA-708

TELECOMMUNICATION INFORMATION:

TELEPHONE: 352-375-8100

TELEFAX: 352-372-5800

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 790 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: 36a

US-08-960-780-4

Query Match 10.5%; Score 66; DB 3; Length 790;

Best Local Similarity 23.4%; Pred. No. 21;

Query Match	10.3%;	Score 65;	DB 4;	Length 587;
Best Local Similarity	23.9%;	Pred. No. 19;		
Matches	21;	Conservative	21;	Mismatches .38;
			Indels	8;
			Gaps	3;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 11:41:40 ; Search time 282.676 Seconds
(without alignments)
84.837 Million cell updates/sec

Title: US-09-890-806-3_COPY_28_143
Perfect score: 629
Sequence: 1 AGEDALRPWKSTAKHPWFQI.....LOGDVIWLRVCELRVDYTS 116

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 777136 seqs, 206736638 residues

Total number of hits satisfying chosen parameters: 777136

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	629	100.0	199	15	US-10-300-393-19
2	91	14.5	186	15	US-10-300-393-20
3	75.5	12.0	301	12	US-10-369-493-2939
4	69	11.0	893	14	US-10-014-436-4
5	69	11.0	1938	14	US-10-014-436-2
6	68	10.8	384	12	US-10-369-493-7954
7	67.5	10.7	448	12	US-10-375-932-25
8	67.5	10.7	464	12	US-10-375-932-89
9	67.5	10.7	564	12	US-10-369-493-585
10	67.5	10.7	6291	12	US-10-329-079-41
11	67	10.7	240	9	US-09-853-161-100
12	67	10.7	240	9	US-09-852-659A-100
13	67	10.7	240	10	US-09-852-797-100
14	67	10.7	335	9	US-09-853-161-64
15	67	10.7	335	9	US-09-852-659A-64

16	67	10.7	335	10	US-09-852-797-64	Sequence 64, Appl
17	67	10.7	824	12	US-10-210-951-58	Sequence 58, Appl
18	67	10.7	824	12	US-10-211-884-58	Sequence 58, Appl
19	67	10.7	824	15	US-10-226-844-1	Sequence 1, Appl
20	66.5	10.6	329	12	US-10-104-047-3709	Sequence 3709, Ap
21	66.5	10.6	449	12	US-10-375-932-38	Sequence 38, Appl
22	66.5	10.6	465	12	US-10-375-932-102	Sequence 102, App
23	66.5	10.6	797	12	US-10-369-493-12480	Sequence 12480, A
24	66.5	10.6	1151	12	US-10-369-493-22104	Sequence 22104, A
25	66.5	10.6	3931	12	US-10-120-801-18	Sequence 18, Appl
26	66	10.5	790	10	US-09-850-351A-4	Sequence 4, Appl
27	65.5	10.4	444	12	US-10-375-932-26	Sequence 26, Appl
28	65.5	10.4	447	12	US-10-375-932-33	Sequence 33, Appl
29	65.5	10.4	460	12	US-10-375-932-90	Sequence 90, Appl
30	65.5	10.4	463	12	US-10-375-932-97	Sequence 97, Appl
31	65.5	10.4	516	9	US-09-883-797-14	Sequence 14, Appl
32	65	10.3	223	12	US-10-137-113-36	Sequence 36, Appl
33	65	10.3	557	12	US-10-369-493-8159	Sequence 8159, Ap
34	64	10.2	469	12	US-10-369-493-18267	Sequence 18267, A
35	64	10.2	789	15	US-10-099-285-78	Sequence 78, Appl
36	64	10.2	789	15	US-10-099-285-80	Sequence 80, Appl
37	64	10.2	789	15	US-10-099-285-92	Sequence 92, Appl
38	64	10.2	789	15	US-10-099-285-94	Sequence 94, Appl
39	64	10.2	789	15	US-10-099-285-96	Sequence 96, Appl
40	64	10.2	790	10	US-09-850-351A-8	Sequence 8, Appl
41	64	10.2	790	12	US-10-452-002A-2	Sequence 2, Appl
42	64	10.2	5245	12	US-10-329-079-45	Sequence 45, Appl
43	63.5	10.1	491	12	US-10-369-493-10822	Sequence 10822, A
44	63	10.0	297	11	US-09-557-796-30	Sequence 30, Appl
45	63	10.0	641	9	US-09-827-040-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-300-393-19
; Sequence 19, Application US/10300393
; Publication No. US20030118568A1
; GENERAL INFORMATION:
; APPLICANT: Crew, Mark D.
; TITLE OF INVENTION: Viral Stealth Technology to Prevent T-Cell-Mediated Rejection of
; FILE REFERENCE: 8793-50391
; CURRENT APPLICATION NUMBER: US/10/300,393
; PRIOR FILING DATE: 2002-11-19
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 199
; TYPE: PRT
; ORGANISM: human cytomegalovirus Towne strain US2
US-10-300-393-19

Query Match 100.0%; Score 629; DB 15; Length 199;
Best Local Similarity 100.0%; Pred. No. 2.6e-68;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGEDALRPWKSTAKHPWFQIEDNRCYIDNGKLFGAGSIVGNNSRFVFPDPAKYGGVGENL 60
DB 28 AGEDALRPWKSTAKHPWFQIEDNRCYIDNGKLFARGSIVGNNSRFVFPDPAKYGGVGENL 87
QY 61 YHADVFEVFGESLKNVNRNLDVMPIFETLRLVLOGDVTLWLRVCELRVDYTS 116
DB 88 YHADVFEVFGESLKNVNRNLDVMPIFETLRLVLOGDVTLWLRVCELRVDYTS 143

RESULT 2
US-10-300-393-20
; Sequence 20, Application US/10300393
; Publication No. US20030118568A1

```

; GENERAL INFORMATION:
; APPLICANT: Crew, Mark D.
; TITLE OF INVENTION: Viral Stealth Technology to Prevent T-Cell-Mediated Rejection of
; FILE REFERENCE: 8793-50391
; CURRENT APPLICATION NUMBER: US/10/300,393
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/342,981
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 186
; TYPE: PRT
; ORGANISM: human cytomegalovirus Towne strain US3
US-10-300-393-20

Query Match          14.5%; Score 91; DB 15; Length 186;
Best Local Similarity 27.4%; Pred. No. 0.0069;
Matches 31; Conservative 20; Mismatches 38; Indels 24; Gaps 6;

QY 18 FOIEDNRCYIDNGKLFARGSGIVGNMSR-----FVFDPKADYGGVGENLYVHADVEFV 70
DB 37 FRVEENQCMHMGMLHYKGRNSGNFTEKHFSVGSVQSYMDRLQVSGEQYHDE----- 91

QY 71 PGESLKNNV-----RNLDPVPIFETLRLVLQGD-VIMLRCPVELRVDTYS 116
DB 92 RGAYFEWNIQGHVPVHTVDWVDI--TLSTR---WGDPKXYAACVQVRMDYSS 139

RESULT 3
US-10-369-493-2939
; Sequence 2939, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2939
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Thermotoga maritima
US-10-369-493-2939

Query Match          12.0%; Score 75.5; DB 12; Length 301;
Best Local Similarity 32.3%; Pred. No. 0.96;
Matches 30; Conservative 11; Mismatches 33; Indels 19; Gaps 5;

QY 19 QIEDNRCYIDNGKLFARG-----SIVGNMSRFVFDPKAD---YGGVGENLYVHADVE 68
DB 205 RLSDHVCILIDHGIAGTSPSSLISSSGIKTVVFDDQVNVNRYLEKKEKHYVETD-- 262

QY 69 FVPGESLK-----WNVNRLDV-MPIFETLRL 95
DB 263 -APBELIKDLKNNVSNIVIRKPNLEDFLKL 294

RESULT 4
US-10-014-436-4
; Sequence 4, Application US/10014436
; Publication No. US20020182699A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, Yuji
```

```

; APPLICANT: IGARASHI, Kazuaki
; APPLICANT: OZAKI, Katsuya
; APPLICANT: ARA, Katsutoshi
; APPLICANT: KAWAI, Shuji
; APPLICANT: ITO, Susumu
; TITLE OF INVENTION: Gene for Enzyme Having Both Alkaline Pullulanase and Alkaline Alpt
; FILE REFERENCE: 2173-0122P
; CURRENT APPLICATION NUMBER: US/10/014,436
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: JP 111547/1995
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: PCT/JP96/01243
; PRIOR FILING DATE: 1996-05-10
; PRIOR APPLICATION NUMBER: US 08/952,084
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: US 09/514,302
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 893
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-10-014-436-4

Query Match          11.0%; Score 69; DB 14; Length 893;
Best Local Similarity 32.0%; Pred. No. 23;
Matches 24; Conservative 11; Mismatches 30; Indels 10; Gaps 3;

QY 12 TAKHPWFQIEDNRCYIDNGKLFARGSGIVGNMSRFVFDPKAD-----YGGVGENLYVHAD 65
DB 82 SVKSGMKLIDENYAY--DKLGAELHEDGTATLTKWSPKADNVSVLYKVDQNEVY--D 137

QY 66 DVEFVPGESLKNNVR 80
DB 138 TIEMVKGDRGVMSVK 152

RESULT 5
US-10-014-436-2
; Sequence 2, Application US/10014436
; Publication No. US20020182699A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, Yuji
; APPLICANT: IGARASHI, Kazuaki
; APPLICANT: OZAKI, Katsuya
; APPLICANT: ARA, Katsutoshi
; APPLICANT: KAWAI, Shuji
; APPLICANT: ITO, Susumu
; TITLE OF INVENTION: Gene for Enzyme Having Both Alkaline Pullulanase and Alkaline Alpt
; FILE REFERENCE: 2173-0122P
; CURRENT APPLICATION NUMBER: US/10/014,436
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: JP 111547/1995
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: PCT/JP96/01243
; PRIOR FILING DATE: 1996-05-10
; PRIOR APPLICATION NUMBER: US 08/952,084
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: US 09/514,302
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1938
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-10-014-436-2

Query Match          11.0%; Score 69; DB 14; Length 1938;
Best Local Similarity 32.0%; Pred. No. 59;
```

Matches 24; Conservative 11; Mismatches 30; Indels 10; Gaps 3;

QY 12 TAKHPWFQIEDNRCYIDNGKLFARGSIVGNMSRFVDPKAD-----YGGVGENLYVHAD 65

Db 1127 SVKSGMKLIDEMKAY--DGKLGAEIHEDGTATLKWSPKADNVSVLVDKVDQNEVY--D 1182

QY 66 DVEFVPGESLKMNV 80

Db 1183 TIEMVKGDRGVMSVK 1197

RESULT 6

US-10-369-493-7954

; Sequence 7954, Application US/10369493
; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 7954

; LENGTH: 384

; TYPE: PRT

; ORGANISM: Rhodobacter sphaeroides

; US-10-369-493-7954

Query Match 10.8%; Score 68; DB 12; Length 384;

Best Local Similarity 27.7%; Pred. No. 11;

Matches 23; Conservative 9; Mismatches 25; Indels 26; Gaps 3;

QY 17 WFOIEDNRCYIDNGKLFARGSIVGNMSRFVDPKADYGGVGENLYVHAD--VEFVPGES 74

Db 19 WLTAEDESRVLDG-----AGIAVNALGHANPALVEVISAQA 55

QY 75 LK-WNVVRLDVMPIFETLALRV 96

Db 56 AKLMHVSNTLYRIPEGERLAEKLV 78

RESULT 7

US-10-375-932-25

; Sequence 25, Application US/10375932
; Publication No. US20040009469A1

; GENERAL INFORMATION:

; APPLICANT: Apt, Doris

; APPLICANT: Punnonen, Juha

; APPLICANT: Brinkman, Alice M.

; TITLE OF INVENTION: NOVEL FLAVIVIRUS ANTIGENS

; FILE REFERENCE: 0322.210US

; CURRENT APPLICATION NUMBER: US/10/375,932

; CURRENT FILING DATE: 2003-02-26

; PRIOR APPLICATION NUMBER: US 60/360,030

; PRIOR FILING DATE: 2002-02-26

; SOFTWARE: FastSeq for windows Version 4.0

; SEQ ID NO 25

; LENGTH: 448

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic polypeptide: 16B4-NPRM

; US-10-375-932-25

Query Match 10.7%; Score 67.5; DB 12; Length 448;

Best Local Similarity 25.0%; Pred. No. 15;

Matches 28; Conservative 12; Mismatches 43; Indels 29; Gaps 6;

QY 2 GEDALRPWKSTAKHPWFQIEDNRCYIDNGKLFARGSIVGNMSRFV-FDPKADY----- 54

Db 327 GEDA-----PCKIP-FSTEDEKVTQNGRLITANPIVTEKDSPVNIEAPFGESEYIV 379

QY 55 GVENLYVHADVEFVPGESLKMNVRLDVMPIFETL---ALRLVLOQDVIV 103

Db 380 GVGDKAL-----KINWYKKGSSIGQMFTTMRGAKRMALIGDTAW 419

RESULT 8

US-10-375-932-89

; Sequence 89, Application US/10375932
; Publication No. US20040009469A1

; GENERAL INFORMATION:

; APPLICANT: Apt, Doris

; APPLICANT: Punnonen, Juha

; APPLICANT: Brinkman, Alice M.

; TITLE OF INVENTION: NOVEL FLAVIVIRUS ANTIGENS

; FILE REFERENCE: 0322.210US

; CURRENT APPLICATION NUMBER: US/10/375,932

; CURRENT FILING DATE: 2003-02-26

; PRIOR APPLICATION NUMBER: US 60/360,030

; PRIOR FILING DATE: 2002-02-26

; NUMBER OF SEQ ID NOS: 345

; SOFTWARE: FastSeq for windows Version 4.0

; SEQ ID NO 89

; LENGTH: 464

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic polypeptide: 16B4

; US-10-375-932-89

Query Match 10.7%; Score 67.5; DB 12; Length 464;

Best Local Similarity 25.0%; Pred. No. 15;

Matches 28; Conservative 12; Mismatches 43; Indels 29; Gaps 6;

QY 2 GEDALRPWKSTAKHPWFQIEDNRCYIDNGKLFARGSIVGNMSRFV-FDPKADY----- 54

Db 343 GEDA-----PCKIP-FSTEDEKVTQNGRLITANPIVTEKDSPVNIEAPFGESEYIV 395

QY 55 GVENLYVHADVEFVPGESLKMNVRLDVMPIFETL---ALRLVLOQDVIV 103

Db 396 GVGDKAL-----KINWYKKGSSIGQMFTTMRGAKRMALIGDTAW 435

RESULT 9

US-10-369-493-585

; Sequence 585, Application US/10369493
; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 585

; LENGTH: 564

; TYPE: PRT

; ORGANISM: Deinococcus radiodurans

; US-10-369-493-585

	Query Match	Best Local Similarity	Matches	Score	Pred.	No. 20;	DB	Length	Gaps
		10.7%;	26.4%;	67.5;	14;	47;	12;	564;	
		Conservative	Matches						
OY	8	PMKSTAKHPWFQIEDNRCYIDNGKLPAR-----GSIVGNMGRFVFPDPKADYCG	55						
Db	136	PNHTSSDHAMFQ-----EALTGKASAKRDWVWRDPAADGGLPNNWKSFFGGEAWTLDE	189						
OY	56	VGENLYHADVEFVPGE-SLKMNVRNLDV-MPIFETLA-LRLVLQG--DYIWLRCVP	108						
Db	190	ASGGYYLH----QFLPSQPDLNW--RNPDRVAAAFDYLRFMRGRGVDDGFVDITWLLAED	243						
OY	109	ELRVD	113						
Db	244	ERFLD	248						

```

RESULT 10
US-10-329-079-41
; Sequence 41, Application US/10329079
; Publication NO. US20030198981A1
; GENERAL INFORMATION:
; APPLICANT: FARNET, Chris
; APPLICANT: ZAZOPOULOS, Emmanuel
; APPLICANT: STAFFA, Alfredo
; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF LIPOPEPTIDES
; FILE REFERENCE: 3002-11US
; CURRENT APPLICATION NUMBER: US/10/329, 079
; CURRENT FILING DATE: 2002-12-24
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 6291
; TYPE: PRF
; ORGANISM: Streptomyces refuineus
US-10-329-079-41

```

Query Match	10.7%;	Score 67.5;	DB 12;	Length 6291;
Best Local Similarity	30.0%;	Pred. No. 3.9e+02;		
Matches	24;	Conservative 7;	Mismatches 22;	Indels 27; Gaps 3;

QY	3	EDALRPWKSTAKHPWFQIEDNRCYIDNGKLPARGSIYG-----NMSRFVDPKADY	53
		: : :	: : :
DB	781	DGALRFPVPDGA-----TGLYLRGGRGLARGYLRRPGLTARFVADP---H	822
		: :	: : :
QY	54	GGVGENLIYHADVEFVPGE	73
		: :	: : :
DB	823	TGTGERMYRTGDLVRRVPGE	842

```

RESULT 11
US-09-853-161-100
; Sequence 100, Application US/09853161
; Patent No. US20020076756A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P3
; CURRENT APPLICATION NUMBER: US/09/853,161
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100

```

```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-853-161-100

```

Query Match	10.7%	Score 67	DB 9	Length 240
Best Local Similarity	26.4%	Pred. No. 7.8		
Matches	29	Conservative 18	Mismatches 35	Indels 28
			Gaps	6
Qy	14	KHPWFQIED-----NRCYIDNGKLFARGSIGVGNMRFVDPKADYGGVGEN-----	59	
Db	37	RHAIVQAEBHLQTAGTCGVSDSL---GSLGPRTAAVFRPR-----GDSLPRRETRYV	88	
Qy	60	-LYVHADDDVER-VPGESLKMNVNLDVM---PIFETLALRLVLOGDIVM	103	
Db	89	ELYYVVDNAEFOMLGSEAAVRHRVLEVVNVHVDKLYOKLNRVVVLGLEIWM	138	

```

US-09-852-659A-100
; Sequence 100, Application US/09852659A
; Patent No. US20020077287A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P4
; CURRENT APPLICATION NUMBER: US/09/852,659A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-852-659A-100

```

```

Query Matchc 10.7%; Score 67; DB 9; Length 240;
Best Local Similarity 26.4%; Pred. No. 7.8;
Matches 29; Conservative 18; Mismatches 35; Indels 28; Gaps 6;

Qy 14 KHPWFQIED-----NRCYIDNGKLFARGSLVGNMGRFVFPDKADYGGVGEN----- 59
Db 37 RHAAVYQAEHLIQTAGTCGVSDDSL--GSLGPRTAAVFRPRP-----GDSLPSRETRYV 88
Qy 60 -LYVHADDDVEF-VPGESLKMVNRNDVM----PIFETLALRLVLQGDVIW 103
Db 89 ELVVVDNAEFQWLGSBAAVRHRVLEVVNHVDKLYQKLNFRVVLVGLIWI 138

RESULT 13
US-09-852-797-100
; Sequence 100, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-852-797-100

Query Matchc 10.7%; Score 67; DB 10; Length 240;
Best Local Similarity 26.4%; Pred. No. 7.8;
Matches 29; Conservative 18; Mismatches 35; Indels 28; Gaps 6;

Qy 14 KHPWFQIED-----NRCYIDNGKLFARGSLVGNMGRFVFPDKADYGGVGEN----- 59
Db 37 RHAAVYQAEHLIQTAGTCGVSDDSL--GSLGPRTAAVFRPRP-----GDSLPSRETRYV 88
Qy 60 -LYVHADDDVEF-VPGESLKMVNRNDVM----PIFETLALRLVLQGDVIW 103
Db 89 ELVVVDNAEFQWLGSBAAVRHRVLEVVNHVDKLYQKLNFRVVLVGLIWI 138

RESULT 14
US-09-853-161-64
; Sequence 64, Application US/09853161
; Patent No. US20020076756A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P3
; CURRENT APPLICATION NUMBER: US/09/853,161
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (35)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (297)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-853-161-64

Query Match          10.7%; Score 67; DB 9; Length 335;
Best Local Similarity 26.4%; Pred.No.12;
Matches 29; Conservative 18; Mismatches 35; Indels 28; Gaps 6

Qy      14 KHPWEQIED----NRCYIDNGKLFARGSIVGNNRFFVDPKADYGGVGEN-----59
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      151 RHAIVQAHEHLIQTAGTIGVSDSL--GSLLGPTAAVFRPR-----GDSLPSRETRYV 202
       ||::||::||::||::||::||::||::||::||::||::||::||::||:
Qy      60 -LYHADDER-VPGESLKMVNRLDVM---PIFETLARLVLOGDVIW 103
       |||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      203 ELVVVDNAERQMGLGEAAVRHRYLEVNVNHVDKLYQKLNRRVLVLGLEIWM 252
       |||::||::||::||::||::||::||::||::||::||::||::||::||:

RESULT 15
US-09-852-659A-64
; Sequence 64, Application US/09852659A
; Patent No. US2002007287A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P4
; CURRENT APPLICATION NUMBER: US/09/852,659A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
```

```

PRIORITY FILING DATE: 1998-03-12
PRIORITY APPLICATION NUMBER: 60/040,762
PRIORITY FILING DATE: 1997-03-14
PRIORITY APPLICATION NUMBER: 60/040,710
PRIORITY FILING DATE: 1997-03-14
PRIORITY APPLICATION NUMBER: 60/050,934
PRIORITY FILING DATE: 1997-05-30
PRIORITY APPLICATION NUMBER: 60/048,100
PRIORITY FILING DATE: 1997-05-30
PRIORITY APPLICATION NUMBER: 60/048,357
PRIORITY FILING DATE: 1997-05-30
PRIORITY APPLICATION NUMBER: 60/048,189
PRIORITY FILING DATE: 1997-05-30
PRIORITY APPLICATION NUMBER: 60/057,765
PRIORITY FILING DATE: 1997-09-05
PRIORITY APPLICATION NUMBER: 60/048,970
PRIORITY FILING DATE: 1997-06-06
PRIORITY APPLICATION NUMBER: 60/068,368
PRIORITY FILING DATE: 1997-12-19
NUMBER OF SEQ ID NOS: 121
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 64
LENGTH: 335
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (35)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (297)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-852-659A-64

```

[illegible]

Search completed: January 22, 2004, 12:09:28
Job time : 283.676 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 11:37:25 ; Search time 21.3788 Seconds
(without alignments)
521.804 Million cell updates/sec

Title: US-09-890-806-3_COPY_28_143
Perfect score: 629

Sequence: 1 AGEDALRPWKSTAKHPWFQI.....LOGDVIWLRCPVRLVDYTS 116

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	629	100.0	199	1	QOBECS	HQLF2 protein prec
2	97	15.4	186	1	QOBECS	HQLF1 precursor -
3	78.5	12.5	254	2	AD0205	probable N-acetyl m
4	78.5	12.5	415	2	T21532	hypothetical prote
5	77.5	12.3	703	1	ALBSX1	cyclomaltoextrin
6	75.5	12.0	301	2	E72384	ABC transporter, A
7	74	11.8	261	2	A98950	hypothetical prote
8	74	11.8	261	2	E85798	hypothetical prote
9	74	11.8	261	2	C64948	probable membrane
10	74	11.8	360	2	A18854	conserved hypothet
11	74	11.8	360	2	F97631	hypothetical prote
12	72.5	11.5	159	1	E71208	hypothetical prote
13	72	11.4	1409	1	OFFFCP	copia polyprotein
14	71	11.3	2201	2	AH0095	probable sideropho
15	70.5	11.2	394	2	AF2963	phosphomannose iso
16	70.5	11.2	394	2	G98319	mannose-6-phosphat
17	70.5	11.2	704	2	I39805	cyclomaltoextrin
18	70	11.1	271	2	C72207	hypothetical prote
19	70	11.1	588	2	T48009	PECTINESTERASE-11k
20	69.5	11.0	349	1	QOBECS	HHRF5 protein - hu
21	68.5	10.9	273	2	D81878	probable oxidoredu
22	68.5	10.9	314	2	D97538	ribosomal large ch
23	68.5	10.9	329	2	AG2757	hypothetical prote
24	68.5	10.9	347	2	D83605	probable binding p
25	68.5	10.9	359	2	A10140	molybdenum transpo
26	68.5	10.9	1629	2	T06461	DNA-binding protei
27	68	10.8	689	2	F81286	probable polysacch
28	68	10.8	1554	2	C72647	hypothetical prote
29	67.5	10.7	564	2	H75403	glycosyl hydrolase

30	67	10.7	367	2	T09376	hypothetical prote
31	67	10.7	381	2	S69795	UTP-hexose-1-phosp
32	67	10.7	757	2	B90572	lipoprotein (impor
33	67	10.7	1690	2	T40847	probable RNA biog
34	67	10.7	7463	2	T36248	CDA peptide synthe
35	66.5	10.6	159	2	F75021	ribosomal protein
36	66.5	10.6	233	2	E86340	protein F2D10.32 l
37	66.5	10.6	463	2	D41853	hexose phosphate t
38	66.5	10.6	463	2	AC0964	hexosephosphate tr
39	66.5	10.6	1151	2	S48431	probable membrane
40	66.5	10.6	1208	2	T27822	hypothetical prote
41	66	10.5	261	2	A10742	high-affinity zinc
42	66	10.5	361	2	JN0716	glutamate-ammonia
43	65.5	10.4	273	2	H81140	oxidoreductase, sh
44	65.5	10.4	286	2	T36880	probable fo-fo par
45	65.5	10.4	489	2	G91143	ribonuclease G (im

ALIGNMENTS

RESULT 1

QOBECS
HQLF2 protein precursor - human cytomegalovirus (strain AD169)

N/Alternate names: hypothetical protein US2

C/Species: human cytomegalovirus, human herpesvirus 5

C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 28-Jul-2000

C/Accession: E26078; S09916

R/Weston, K.; Barrell, B.G.

J. Mol. Biol. 192, 177-208 (1986)

A/Title: Sequence of the short unique region, short repeats, and part of the long repeat

A/Reference number: A92935; MUID:87169717; PMID:3031311

A/Accession: E26078

A/Molecule type: DNA

A/Residues: 1-199 <MES>

A/Cross-references: EMBL:X04650; NID:G59801; PIDN:CAB37096.1; PID:G4456177

A/Experimental source: strain AD169

R/Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Hornell, T.; M.

M.; Barrell, B.G.

Curr. Top. Microbiol. Immunol. 154, 125-169, (1990)

A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus

A/Reference number: S09749; MUID:90289039; PMID:2161319

A/Accession: S09916

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-199 <CHE>

A/Cross-references: EMBL:X17403; NID:G59591; PIDN:CA35313.1; PID:G1780933

A/Experimental source: strain AD169

A/Note: this sequence was submitted to the EMBL Data Library, December 1989

C/Genetics:

C/Superfamily: cytomegalovirus HQLF2 protein

C/Keywords: glycoprotein; transmembrane protein

F/1-25/Domain: signal sequence #status predicted <SIG>

F/26-199/Product: hypothetical protein US2 #status predicted <MAT>

F/161-186/Domain: transmembrane #status predicted <TMM>

F/68/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 629; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 5.8e-58;
Matches 116; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AGEDALRPWKSTAKHPWFQIEDNRCYIDNGKLFARGSVGNNSRFVDPKADYGGVGENL	60
DB	28	AGEDALRPWKSTAKHPWFQIEDNRCYIDNGKLFARGSVGNNSRFVDPKADYGGVGENL	87
QY	61	YVHADVEFVGESLKNVRLDVMPIFETLALRLVLOGDVWLRCPVRLVDYTS	116
DB	88	YVHADVEFVGESLKNVRLDVMPIFETLALRLVLOGDVWLRCPVRLVDYTS	143

RESULT 2
QOBECS

```

HQLF1 precursor - human cytomegalovirus (strain AD169)
N:Alternate names: hypothetical protein US3
C:Species: human cytomegalovirus, human herpesvirus 5
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 28-Jul-2000
C:Accession: F26078; S09917
R:Weston, K.; Barrell, B.G.
J. Mol. Biol. 192, 177-208, 1986
A:Title: Sequence of the short unique region, short repeats, and part of the long repeat
A:Reference number: A92935; MUID:87169717; PMID:3031311
A:Accession: F26078
A:Molecule type: DNA
A:Residues: 1-186 <WES>
A:Cross-references: EMBL:X04650; NID:g559801; PIDN:CAB37097.1; PID:g4456178
A:Experimental source: strain AD169
R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Hornsall, T.;
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A:Reference number: S09749; MUID:90269039; PMID:2161319
A:Accession: S09917
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-186 <CHR>
A:Cross-references: EMBL:X17403; NID:g59591; PIDN:CAA35314.1; PID:g1780934
A:Experimental source: strain AD169
A:Note: this sequence was submitted to the EMBL Data Library, December 1989
C:Genetics:
A:Gene: HQLF1
C:Superfamily: cytomegalovirus HQLF1 protein
C:Keywords: glycoprotein; transmembrane protein
F:1-25/Domain: signal sequence #status predicted <Sig>
F:26-186/Product: hypothetical protein US3 #status predicted <MAT>
F:160-182/Domain: transmembrane #status predicted <TM>
F:60/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match          15.4%; Score 97; DB 1; Length 186;
Best Local Similarity 27.0%; Pred. No. 0.01;
Matches 31; Conservative 24; Mismatches 32; Indels 28; Gaps 7;

Qy      18 FOIEDNRXYIDNGKLFARGSIGVNMGRFVPDPKADYGGVGNL-----YHADVE 68
      ||::||:| | | | | | | | | | | | | | | | | | | | | | | | | |
Db      37 FRVENQCFHFMGMLEFKGNMSGNFTEKHF--VNVGISQSYMDRLQVSGEQYHDE-- 91

Qy      69 FVPGESLKNV-----RNDVMPFETLALRLVLSQD-VIMLRCPPELRVDYTS 116
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      92 --RGAYFEMNIGGHPVTHVTVDMDVI--TLSTR---WGDPKCYAACVPQVMDYSS 139

RESULT 3
AD0205
probable N-acetylmuramoyl-L-alanine amidase [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AD0205
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AD0205
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-254 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC90504.1; PID:g15979715; GSPDB:GN00175
C:Genetics:
A:Gene: YPO1683

Query Match          12.5%; Score 78.5; DB 2; Length 254;
Best Local Similarity 25.3%; Pred. No. 1.2;
Matches 24; Conservative 15; Mismatches 43; Indels 13; Gaps 3;

8 PWKS--TAKHPWFOIEDNRXYIDN--GKLFARGSIGVNMGRFVPDPKADYGGVGNLTV 62

```

[illegible]

A;Cross-references: FlyBase:FBgn0000349
A;Mobile element: retrotransposon copia
C;Superfamily: retrovirus-related polyprotein
C;Keywords: polyprotein; proteinase
F;2-433/Product: copia protein, 48k #status predicted <MAT1>
F;2-270/Product: copia protein, 31k #status predicted <MAT2>
F;271-433/Product: proteinase #status predicted <MAT3>

Query Match	11.4%;	Score 72;	DB 1;	Length 1409;
Best Local Similarity	23.5%;	Pred. No. 45;		
Matches	24;	Conservative	18;	Mismatches 34;
				Indels 26;
				Gaps 5;

```

0Y      17 WFOIED--NRCTIDNGKLFARGSIVGNNSRFVPEPKADYGVGENTLV--HADDVEFVP 71
      ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db     1048 WFEVEFQALKECEFN-----SSVDRCTY--ILDKKNINENYVLLYVDVVIAT 1095

```

```
QY      72 GESLKW-----VRNDVMPFEETALRLVLQGDVIML 104
      | : | | : | : : | : | : |
Db      1096 GDMTRMNNFRKRYLMEKFRMTDLNELKHFIGIRILEWQDKITYL 1137
```

RESULT 14

AH0095

C:species: *Yersinia pestis*
C:date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001

R. Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B. deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001

A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague
A:Reference number: AB00001; MUID:21470413; PMID:11586360

A;Status: preliminary
A;molecule type: DNA

A;Residues: 1-2201 <KUR>

A; Cross-references: GB:AL590842; PIDN:CAC89627.1; PID:g15978856; GSPDB:GN001755

C; Genetics:

Query Match	Score	DB	Length
11.3%	71	2	2201

Best Local Similarity 28.2%; Pred. No. 97;
Matches 33; Conservative 24; Mismatches 26; Indels 34; Gaps 8;

```
OY      21 EDNRCYIDNGK--LFARGSLVGNMSRF-VFDPADYGCGENLYI-HADVEFVPGSLK 76
       |||:::||||:::||:::||:::||:::||:::||:::||:::||:::||:::||
Db     911 EINRRYDDGKITLFAISALNFDSLTYDIFGPIS----VGGLVLPNAGDEK---EAKQ 965
```

```

0y      77 W-----NVRNLDVMP-IFETLAL-----RLVIQGDVIMRCVPEIR 111
        | : : : | : : : | : : : | : : : |
Db      963 WLSALHQQVTHMNSVPALFEMLLIAEGTQALPRSLQGVLLSGDWITGLDLLPRR 1019

```

RESULT 15

AF2963

C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence 11-Jan-2002 #text_change 18-Nov-2002

R. Wood, D. W.; Setubal, J. C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G. E.; Chen, Y.; Woo, I. E.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCellnell, J.; Karp, P.; Romero, P.; Zhang, S.

Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens
A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AF2963
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-394 <KUR>

A;Cross-references: GB:AE008689; PIDN:AAL44124.1; PID:g17741695; GSPDB:GN00187

A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu3311
A;Map position: linear chromosome

Query Match	11.2%;	Score 70.5;	DB 2;	Length 394;
Best Local Similarity	27.1%;	Pred. No. 14;		
Matches 23;	Conservative 12;	Mismatches 39;	Indels 11;	Gaps 5;

QY 8 PMKSTAKH--PMFQIEDNRCY-IDNGKLFPAGSIVGNMSRFVDPKADYGSGVENLYVHA 64
||| ||: : ||: ||: ||: ||:
Db 82 PMKDASHGLSWFE---KYRLENG--LYGNLADQTGRLL-DPSFDLYNQAFALFAAA 133

Qy 65 DDVEFVPGESLKNVNRNDVMPIFE 89
: : : : :
Db 134 QAASLPERKDEMRVRAIDLALTE 158

Search completed: January 22, 2004, 11:42:39
Job time : 23.3788 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 11:41:40 ; Search time 431.324 Seconds
(without alignments)
84.837 Million cell updates/sec

Title: US-09-890-806-5
Perfect score: 947
Sequence: 1 MKFLVNLVFMVVISYIY.....GQELAPDEPDSALLDPV 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 777136 seqs, 206736638 residues
Total number of hits satisfying chosen parameters: 777136

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	740	78.1	199	15	US-10-300-393-19 Sequence 19, Appl
2	117.5	12.4	451	14	US-10-029-009-20 Sequence 20, Appl
3	115	12.1	186	15	US-10-300-393-20 Sequence 20, Appl
4	110	11.6	498	14	US-10-029-009-22 Sequence 22, Appl
5	109.5	11.6	1189	15	US-10-044-692-325 Sequence 325, App
6	109.5	11.6	1189	15	US-10-044-539-325 Sequence 325, App
7	104	11.0	496	14	US-10-029-009-34 Sequence 34, Appl
8	102	10.8	21	9	US-09-766-378A-36 Sequence 36, Appl
9	102	10.8	21	11	US-09-927-122-3 Sequence 3, Appl
10	102	10.8	21	11	US-09-927-121B-3 Sequence 3, Appl
11	102	10.8	21	12	US-10-361-849-8 Sequence 8, Appl
12	102	10.8	21	12	US-10-029-009-24 Sequence 24, Appl
13	102	10.8	49	14	US-10-256-977-10 Sequence 10, Appl
14	102	10.8	49	15	US-10-084-298-10 Sequence 10, Appl
15	102	10.8	70	11	US-09-847-208-26 Sequence 26, Appl

16	102	10.8	129	15	US-10-152-190-6 Sequence 6, Appl
17	102	10.8	260	15	US-10-264-634-23 Sequence 23, Appl
18	102	10.8	260	15	US-10-264-634-39 Sequence 39, Appl
19	99	10.5	41	12	US-10-410-842A-40 Sequence 40, Appl
20	99	10.5	41	12	US-10-410-842A-44 Sequence 44, Appl
21	99	10.5	61	12	US-10-410-842A-30 Sequence 30, Appl
22	99	10.5	104	12	US-10-410-842A-6 Sequence 6, Appl
23	99	10.5	394	12	US-10-369-300-13 Sequence 13, Appl
24	99	10.5	394	12	US-10-410-842A-2 Sequence 2, Appl
25	97	10.2	20	14	US-10-029-009-12 Sequence 12, Appl
26	97	10.2	21	12	US-10-133-973-74 Sequence 74, Appl
27	91	9.6	41	12	US-10-410-842A-56 Sequence 56, Appl
28	91	9.6	41	12	US-10-410-842A-69 Sequence 69, Appl
29	86	9.1	6291	12	US-10-329-079-41 Sequence 41, Appl
30	82	8.7	893	14	US-10-014-436-4 Sequence 4, Appl
31	82	8.7	1938	14	US-10-014-436-2 Sequence 2, Appl
32	79	8.3	329	12	US-10-104-047-3709 Sequence 3709, Ap
33	79	8.3	3931	12	US-10-120-801-18 Sequence 18, Appl
34	76.5	8.1	316	10	US-09-736-457-340 Sequence 340, App
35	76.5	8.1	316	10	US-09-902-941-340 Sequence 340, App
36	76.5	8.1	316	10	US-09-849-626-340 Sequence 340, App
37	76.5	8.1	316	11	US-09-476-300-340 Sequence 340, App
38	76.5	8.1	316	12	US-10-274-375-3 Sequence 3, Appl
39	76.5	8.1	316	12	US-10-113-872-340 Sequence 340, App
40	76.5	8.1	316	15	US-10-017-754-340 Sequence 340, App
41	76.5	8.1	358	9	US-09-925-299-923 Sequence 923, App
42	76.5	8.1	358	11	US-09-925-299-923 Sequence 923, App
43	75.5	8.0	248	12	US-10-131-409-85 Sequence 85, Appl
44	75.5	8.0	248	12	US-10-150-811-85 Sequence 85, Appl
45	75.5	8.0	267	12	US-10-131-409-83 Sequence 83, Appl

ALIGNMENTS

RESULT 1
US-10-300-393-19
; Sequence 19, Application US/10300393
; Publication No. US20030118568A1
; GENERAL INFORMATION:
; APPLICANT: Crew, Mark D.
; TITLE OF INVENTION: Viral Stealth Technology to Prevent T-Cell-Mediated Rejection of
; TITLE OF INVENTION: Xenografts
; FILE REFERENCE: 8793-50391
; CURRENT APPLICATION NUMBER: US/10/300,393
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/342,981
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 199
; TYPE: PRT
; ORGANISM: human cytomegalovirus Towne strain US2
US-10-300-393-19

Query Match	78.1%;	Score 740;	DB 15;	Length 199;
Best Local Similarity	100.0%;	Pred. No. 1.5e-77;		
Matches 136;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	22	RLPDGTTKAGEDALRPMKSTAKHPFOIEDNRCYIDNGKLFAFGSIVGNMRSFVDPKAD	81	
DB	20	RLPDGTTKAGEDALRPMKSTAKHPFOIEDNRCYIDNGKLFAFGSIVGNMRSFVDPKAD	79	
QY	82	YGVGGENLYVHADVDFVPGESLKMVNRNIDVMPFETTLALRLVLOGDVIWLRCPBELRV	141	
DB	80	YGVGGENLYVHADVDFVPGESLKMVNRNIDVMPFETTLALRLVLOGDVIWLRCPBELRV	139	
QY	142	DYTSAYMMNMQYGMV	157	
DB	140	DYTSAYMMNMQYGMV	155	

RESULT 2
US-10-029-009-20
; Sequence 20, Application US/10029009
; Publication No. US20020164617A1
; GENERAL INFORMATION:
; APPLICANT: Felsch, Jason S.
; APPLICANT: Annis, David Allen
; APPLICANT: Kalghatgi, Krishna
; APPLICANT: Nash, Huw M.
; TITLE OF INVENTION: Affinity Selection-Based Screening of Hydrophobic Proteins
; FILE REFERENCE: 11025.173 US2
; CURRENT APPLICATION NUMBER: US/10/029, 009
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/258,970
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Human
US-10-029-009-20

Query Match 12.4%; Score 117.5; DB 14; Length 451;
Best Local Similarity 27.7%; Pred. No. 8.3e-05;
Matches 41; Conservative 13; Mismatches 45; Indels 49; Gaps 3;

QY 1 MKFLVNVALVFMVVISYIYARLPDGTAKGEDALRPWKSTAKHPWFQIEDNRCYIDNGK 60
DB 1 MKFLVNVALVFMVVISYIYADYKDDDDKMGQP----- 33
QY 61 LPARGSIVNMSRFVFDPKADYGVGENLYHADDFEYVPGESLKNVNRNLDMPIRETL 120
DB 34 -----GNGSAFLIAPNRS-----HAPDHDVTQORDEVAVVG---MGIWMSL 71
QY 121 ALRLVLQGDVIMLRCPPELRDYTSAY 148
DB 72 IVLAIVFGNVLVITAIKFERLQTVTNY 99

RESULT 3
US-10-300-393-20
; Sequence 20, Application US/10300393
; Publication No. US20030118568A1
; GENERAL INFORMATION:
; APPLICANT: Crew, Mark D.
; TITLE OF INVENTION: Viral Stealth Technology to Prevent T-Cell-Mediated Rejection of
; FILE REFERENCE: 8793-50391
; CURRENT APPLICATION NUMBER: US/10/300,393
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/342,981
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 186
; TYPE: PRT
; ORGANISM: human cytomegalovirus Towne strain US3
US-10-300-393-20

Query Match 12.1%; Score 115; DB 15; Length 186;
Best Local Similarity 24.1%; Pred. No. 4.9e-05;
Matches 38; Conservative 32; Mismatches 60; Indels 28; Gaps 7;

QY 10 VFMVVISYIYARLPDGTAKGEDALRPWKSTAKHPWFQIEDNRCYIDNGKLFARGSIVG 69
DB 4 VLVLAIVLVFLRLADSVPRPLDVV-----SEIRSAHFRVENQCSHMGMLHYKGRMSG 59
QY 70 NMGR-----FVFDKADYGVGENLYHADDFEYVPGESLKNV-----RLDVMP 116
DB 60 NFEKEHFVSIGVSGSYMDRLQVSGEYHDE-----RGAYFEWNIGGHPVPHTVDMVDI 114

QY 117 FETLALRLVLQGD-VIMLRCPPELRDYTSAYMNMNQ 153
DB 115 --TLSTR--WGDPKKYACVPQVMDYSHTINWYLO 147

RESULT 4
US-10-029-009-22
; Sequence 22, Application US/10029009
; Publication No. US20020164617A1
; GENERAL INFORMATION:
; APPLICANT: Felsch, Jason S.
; APPLICANT: Annis, David Allen
; APPLICANT: Kalghatgi, Krishna
; APPLICANT: Nash, Huw M.
; TITLE OF INVENTION: Affinity Selection-Based Screening of Hydrophobic Proteins
; FILE REFERENCE: 11025.173 US2
; CURRENT APPLICATION NUMBER: US/10/029, 009
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/258,970
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human
US-10-029-009-22

Query Match 11.6%; Score 110; DB 14; Length 498;
Best Local Similarity 51.8%; Pred. No. 0.00071;
Matches 29; Conservative 0; Mismatches 13; Indels 14; Gaps 2;

QY 1 MKFLVNVALVFMVVISYIYARLPDGTAKGEDA-----LRPWKSTAKHPW 46
DB 1 MKFLVNVALVFMVVISYIYADYKDDDDKQNTSAPPAVSPNTVLAP-----GKGPM 52

RESULT 5
US-10-044-692-325
; Sequence 325, Application US/10044692
; Publication No. US20030096344A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
; THERAPEUTIC METHODS
; NUMBER OF SEQUENCES: 335
; CORRESPONDENCE ADDRESS:
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/044,692
; FILING DATE: 11-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/912,951
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851,843
 FILING DATE: 06-MAY-1997
 APPLICATION NUMBER: US 08/846,017
 FILING DATE: 25-APR-1997
 APPLICATION NUMBER: US 08/844,419
 FILING DATE: 18-APR-1997
 APPLICATION NUMBER: US 08/724,643
 FILING DATE: 01-OCT-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 015389-002600US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 325:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1189 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 325:
 US-10-044-692-325

Query Match 11.6%; Score 109.5; DB 15; Length 1189;
 Best Local Similarity 32.1%; Pred. No. 0.0026;
 Matches 50; Conservative 10; Mismatches 65; Indels 31; Gaps 6;

QY 1 MKFLVVALVWVYISYIYARLPDGTAG--EDA-----LRPWKSTAKHPWQI 49
 1 MKFLVVALVWVYISYIYAD-PSSRSAGTMEFAASTORCVLRTWEALAPATPAMP 59
 DB 50 EDNRCTIDNGKLFARGSIVGNMRFVDFPKADYGVGENLYVHADVEF-----V 99
 60 RAPRCRAVRSLRSHYREVLPLATFV---RLGPQGWRLVQGDPAARFALVAQCLVCV 115
 QY 100 PGESLKMVRNLDVMPFETLRLVLQGDVIWLRC 135
 DB 116 P-----WDARPPAPSPFRQVSCIKELVARVLQRLC 146

RESULT 6
 US-10-044-539-325
 Sequence 325, Application US/10044539
 Publication No. US20030100093A1
 GENERAL INFORMATION:
 APPLICANT: Cecch, Thomas R.
 Lingner, Joachim
 Nakamura, Toru
 Chapman, Karen B.
 Morin, Gregg B.
 Harley, Calvin
 Andrews, William H.
 TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
 THERAPEUTIC METHODS
 NUMBER OF SEQUENCES: 335
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: United States of America
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/044,539
 FILING DATE: 11-Jan-2002
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/912,951
 FILING DATE: <Unknown>
 APPLICATION NUMBER: US 08/854,050
 FILING DATE: 09-MAY-1997
 APPLICATION NUMBER: US 08/851,843
 FILING DATE: 06-MAY-1997
 APPLICATION NUMBER: US 08/846,017
 FILING DATE: 25-APR-1997
 APPLICATION NUMBER: US 08/844,419
 FILING DATE: 18-APR-1997
 APPLICATION NUMBER: US 08/724,643
 FILING DATE: 01-OCT-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 015389-002600US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 325:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1189 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 325:
 US-10-044-539-325

Query Match 11.6%; Score 109.5; DB 15; Length 1189;
 Best Local Similarity 32.1%; Pred. No. 0.0026;
 Matches 50; Conservative 10; Mismatches 65; Indels 31; Gaps 6;

QY 1 MKFLVVALVWVYISYIYARLPDGTAG--EDA-----LRPWKSTAKHPWQI 49
 1 MKFLVVALVWVYISYIYAD-PSSRSAGTMEFAASTORCVLRTWEALAPATPAMP 59
 DB 50 EDNRCTIDNGKLFARGSIVGNMRFVDFPKADYGVGENLYVHADVEF-----V 99
 60 RAPRCRAVRSLRSHYREVLPLATFV---RLGPQGWRLVQGDPAARFALVAQCLVCV 115
 QY 100 PGESLKMVRNLDVMPFETLRLVLQGDVIWLRC 135
 DB 116 P-----WDARPPAPSPFRQVSCIKELVARVLQRLC 146

RESULT 7
 US-10-029-009-34
 Sequence 34, Application US/10029009
 Publication No. US20020164617A1
 GENERAL INFORMATION:
 APPLICANT: Felsch, Jason S.
 Annis, David Allen
 APPLICANT: Kalghatgi, Krishna
 APPLICANT: Nash, Huw M.
 TITLE OF INVENTION: Affinity Selection-Based Screening of Hydrophobic Proteins
 FILE REFERENCE: 111025.173 US2
 CURRENT APPLICATION NUMBER: US/10/029,009
 CURRENT FILING DATE: 2002-03-28
 PRIOR APPLICATION NUMBER: US 60/258,970
 PRIOR FILING DATE: 2000-12-29
 NUMBER OF SEQ ID NOS: 34
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 34
 LENGTH: 496
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Met-Melitin SS-FLAG M1 tag-m2 mACHR sequence
 US-10-029-009-34
 Query Match 11.0%; Score 104; DB 14; Length 496;

Best Local Similarity 100.0%; Pred. No. 8.4e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKFLVNVALVEMVYISYIYA 21
Db 1 MKFLVNVALVEMVYISYIYA 21

RESULT 12

US-10-029-009-24
; Sequence 24, Application US/10029009
; Publication No. US20020164617A1
; GENERAL INFORMATION:
; APPLICANT: Felsch, Jason S.
; APPLICANT: Annis, David Allen
; APPLICANT: Kalgatgi, Krishna
; APPLICANT: Nash, Huw M.
; TITLE OF INVENTION: Affinity Selection-Based Screening of Hydrophobic Proteins
; FILE REFERENCE: 11025.173 US2
; CURRENT APPLICATION NUMBER: US/10/029,009
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/258,970
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Melittin signal sequence
US-10-029-009-24

Query Match 10.8%; Score 102; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.4e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKFLVNVALVEMVYISYIYA 21
Db 1 MKFLVNVALVEMVYISYIYA 21

RESULT 13

US-10-256-977-10
; Sequence 10, Application US/10256977
; Publication No. US20030157106A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: Pittman, Debra
; APPLICANT: Fouser, Lynette
; APPLICANT: Spaulding, Vikki
; APPLICANT: Xuan, Dejun
; TITLE OF INVENTION: Composition and Method for Treating Inflammatory
; FILE REFERENCE: G15358 CIP
; CURRENT APPLICATION NUMBER: US/10/256,977
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US/10/084,298
; PRIOR FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 60/270,823
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/281,353
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/131,473
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/561,811
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Artificial Sequence

FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid tag
US-10-256-977-10

Query Match 10.8%; Score 102; DB 12; Length 49;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKFLVNVALVEMVYISYIYA 21
Db 1 MKFLVNVALVEMVYISYIYA 21

RESULT 14

US-10-084-298-10
; Sequence 10, Application US/10084298
; Publication No. US20030099649A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: Pittman, Debra
; APPLICANT: Fouser, Lynette
; APPLICANT: Spaulding, Vikki
; APPLICANT: Xuan, Dejun
; TITLE OF INVENTION: Composition and Method for Treating Inflammatory
; FILE REFERENCE: G15358 CIP
; CURRENT APPLICATION NUMBER: US/10/084,298
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 60/270,823
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/281,353
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/131,473
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/561,811
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid tag
US-10-084-298-10

Query Match 10.8%; Score 102; DB 15; Length 49;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKFLVNVALVEMVYISYIYA 21
Db 1 MKFLVNVALVEMVYISYIYA 21

RESULT 15

US-09-847-208-26
; Sequence 26, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67,002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Artificial Sequence

; ORGANISM: Apis mellifera(Honeybee)Apis cerana(Ind. honeybee)
US-09-847-208-26

Query Match 10.8%; Score 102; DB 11; Length 70;
Best Local Similarity 100.0%; Pred.No. 0.00043;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLVNVALLVFMVYYSIYA 21
|||||
Db 1 MKFLVNVALLVFMVYYSIYA 21

Search completed: January 22, 2004, 12:09:28
Job time : 431.324 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 11:37:25 ; Search time 33.2253 Seconds
(without alignments)
225.401 Million cell updates/sec

Title: US-09-890-806-5
Perfect score: 947
Sequence: 1 MKFLVNLVFMVYISYIV.....GQPELAPEDPEDSALLDPV 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	109.5	11.6	1189	US-08-974-549A-613	Sequence 613, App
2	109.5	11.6	1189	US-08-912-951-325	Sequence 325, App
3	102	10.8	21	US-08-960-190A-36	Sequence 36, Appl
4	102	10.8	21	US-09-376-330-4	Sequence 4, Appl1
5	99	10.5	308	US-08-499-568-2	Sequence 2, Appl1
6	99	10.5	308	US-08-793-958-2	Sequence 2, Appl1
7	99	10.5	369	US-08-139-609-1	Sequence 1, Appl1
8	99	10.5	383	5470718-5	Patent No. 5470718
9	99	10.5	388	US-08-499-568-11	Sequence 11, Appl
10	99	10.5	388	US-08-793-958-11	Sequence 11, Appl
11	99	10.5	333	US-08-220-151-23	Sequence 23, Appl
12	99	10.5	333	US-08-413-118-23	Sequence 23, Appl
13	99	10.5	393	US-08-473-446-23	Sequence 23, Appl
14	99	10.5	393	5182195-10	Patent No. 5182195
15	99	10.5	394	US-08-499-568-4	Sequence 4, Appl1
16	99	10.5	394	US-08-793-958-4	Sequence 4, Appl1
17	97	10.2	21	US-08-334-669-2	Sequence 2, Appl1
18	97	10.2	21	US-08-485-341A-2	Sequence 2, Appl1
19	91	9.6	368	PCT-US93-11703-24	Sequence 24, Appl
20	91	9.6	333	US-08-499-568-15	Sequence 15, Appl
21	91	9.6	333	US-08-793-958-15	Sequence 15, Appl
22	91	9.6	393	US-08-956-998-2	Sequence 2, Appl1
23	87.5	9.2	26	US-08-782-480-43	Sequence 43, Appl
24	87.5	9.2	26	US-08-954-211-43	Sequence 43, Appl
25	87.5	9.2	26	US-09-005-167A-43	Sequence 43, Appl
26	87.5	9.2	26	US-09-176-741B-43	Sequence 43, Appl
27	82	8.7	893	US-09-514-302-4	Sequence 4, Appl1

28	82	8.7	1938	4	US-09-514-302-2	Sequence 2, Appl1
29	80	8.4	676	3	US-08-947-965-71	Sequence 71, Appl
30	78	8.2	310	4	US-09-252-991A-17381	Sequence 17381, A
31	77.5	8.2	729	4	US-09-252-991A-32535	Sequence 32535, A
32	76.5	8.1	316	4	US-09-702-705-340	Sequence 340, App
33	76.5	8.1	316	4	US-09-736-457-340	Sequence 340, App
34	72	7.6	675	3	US-08-947-965-76	Sequence 76, Appl
35	71.5	7.6	1266	4	US-08-506-296B-4	Sequence 4, Appl1
36	71	7.5	405	4	US-09-252-991A-28582	Sequence 28582, A
37	70.5	7.4	327	4	US-09-107-532A-6181	Sequence 6181, Ap
38	70	7.4	448	4	US-09-724-623-84	Sequence 84, Appl
39	69.5	7.3	358	1	US-08-604-913B-11	Sequence 11, Appl
40	69.5	7.3	472	4	US-08-985-492-13	Sequence 13, Appl
41	69.5	7.3	521	1	US-08-276-213-3	Sequence 3, Appl1
42	68.5	7.2	130	4	US-08-826-134-12	Sequence 12, Appl
43	68.5	7.2	640	4	US-09-252-991A-27542	Sequence 27542, A
44	68.5	7.2	1384	4	US-08-826-134-2	Sequence 2, Appl1
45	68	7.2	316	1	US-08-585-595-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-974-549A-613
; Sequence 613, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/915,503
 FILING DATE: 14-AUG-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US97/17618
 FILING DATE: 01-OCT-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US97/17885
 FILING DATE: 01-OCT-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph Ted
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 015389-002610US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 613:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1189 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 1..1189
 OTHER INFORMATION:
 OTHER INFORMATION: /note= "fusion protein composed of
 OTHER INFORMATION: hTERT protein"

```

Query Match          11.6%; Score 109.5; DB 3; Length 1189;
Best Local Similarity 32.1%; Pred. No. 0.00063;
Matches 50; Conservative 10; Mismatches 65; Indels 31; Gaps 6;

QY      1 MKFLVVALVFMVVYISYIYARLPDGTGKAG--EDA-----LRPKSTAKHPFQI 49
DB      1 MKFLVVALVFMVVYISYIYAD--SSRSAGTMEFAASTQRCLLTWEALAPATPAMP 59
QY      50 EDNRICYDNGKLFARGSIYGNMSRFVFDPRADYGGVGENLYHADVEF-----V 99
DB      60 RAPRCRAVRSLSRSHYREVLPATFV---RRLGPGWRVLVQRGDPAPAFRALVAQCLYCV 115
QY      100 PGESLKNNVRNLDVMPITFETLALRLVLQGDYIWLRC 135
DB      116 P-----WDARPPAPAPSPFRQVSCLELVARVLQRLC 146

RESULT 2
US-08-912-951-325
; Sequence 325, Application US/08912951
; Patent No. 6475789
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
; NUMBER OF SEQUENCES: 335
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/912,951
FILING DATE: 14-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002600US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 325:
SEQUENCE CHARACTERISTICS:
LENGTH: 1189 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-912-951-325

Query Match      11.6%; Score 109.5; DB 4; Length 1189;
Best Local Similarity 32.1%; Pred. No. 0.00063;
Matches 50; Conservative 10; Mismatches 65; Indels 31; Gaps 6;

QY      1 MKFLVNVAVLVEMVVYISYIYARLPDGIITKAG--EDA-----LRPWKSTAKHPWFQI 49
      |||||
DB      1 MKFLVNVAVLVEMVVYISYIYAD-PSSRSAAGTMEFPAASTQRCVLIRTWELALPATPAMP 59
      |||||

QY      50 EDNRCTIDNGKLFARGSIGVNNRSFVDPKADYGVGGENLYVHADVEF-----V 99
      |||
DB      60 RAPRCRAVRSLSLRSHKREVLPLATFV----RRLLPQGWRLVQRGDPAAFRLVAQCLVCV 115
      |||

QY      100 PGESLKNVNRNDVMPITFETLALRLVLQGDVYIWLRC 135
      |::|
DB      116 P-----WDARPPPAAPSFRRQVSLKELVARVLQRLC 146
      |::|

RESULT 3
US-08-960-190A-36
; Sequence 36, Application US/08960190A
; Patent No. 6232445
; GENERAL INFORMATION:
; APPLICANT: Rhode, Peter R.
; APPLICANT: Acevedo, Jorge
; APPLICANT: Burkhardt, Martin
; APPLICANT: Jiao, Jin-an
; APPLICANT: Wong, Hing C.
; TITLE OF INVENTION: SOLUBLE MHC COMPLEXES AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 38

```

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: usa
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960,190A
FILING DATE: 29-OCT-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-960-190A-36

Query Match 10.8%; Score 102; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKFLVNVALVFMVYISYIYA 21
Db 1 MKFLVNVALVFMVYISYIYA 21

RESULT 4

US-09-376-330-4
Sequence 4, Application US/09376330
Patent No. 6399321

GENERAL INFORMATION:
APPLICANT: Tessier, Daniel C.
APPLICANT: Dignard, Daniel
APPLICANT: Bergeron, John J.M.
APPLICANT: Thomas, David Y.
TITLE OF INVENTION: Method for screening for
TITLE OF INVENTION: UDP-glucose:glycoprotein glucosyltransferase (UGGT) activity
TITLE OF INVENTION: and nucleic acid encoding for UGGT
FILE REFERENCE: 2139-9"US"
CURRENT APPLICATION NUMBER: US/09/376,330
CURRENT FILING DATE: 1999-08-18
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 21
TYPE: PRT
ORGANISM: Melittin signal peptide sequence
US-09-376-330-4

Query Match 10.8%; Score 102; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKFLVNVALVFMVYISYIYA 21

Db 1 MKFLVNVALVFMVYISYIYA 21

RESULT 5

US-08-499-568-2
Sequence 2, Application US/08499568
Patent No. 5654174

GENERAL INFORMATION:
APPLICANT: Cohen, Gary H.
APPLICANT: Eisenberg, Roselyn J.
APPLICANT: Nicola, Anthony
TITLE OF INVENTION: Herpes Simplex Virus Glycoprotein D
TITLE OF INVENTION: Variants
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/499,568
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 5654174and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32813
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-499-568-2

Query Match 10.5%; Score 99; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 159 QPELAPEDPDSALLDPV 177
Db 267 QPELAPEDPDSALLDPV 285

RESULT 6

US-08-793-958-2
Sequence 2, Application US/08793958
Patent No. 5814486

GENERAL INFORMATION:
APPLICANT: Cohen, Gary H.
APPLICANT: Eisenberg, Roselyn J.
APPLICANT: Nicola, Anthony
TITLE OF INVENTION: Herpes Simplex Virus Glycoprotein D
TITLE OF INVENTION: Variants
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA

Query Match 10.8%; Score 102; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKFLVNVALVFMVYISYIYA 21

ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,958
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/499,568
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 581448and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32813
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-793-958-2

Query Match 10.5%; Score 99; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLLEDPV 177
DB 267 QPELAPEDPEDSALLLEDPV 285

RESULT 7
US-08-139-609-1
Sequence 1, Application US/08139609
Patent No. 5837249
GENERAL INFORMATION:
APPLICANT: Heber-Katz, Ellen
APPLICANT: Dietzschold, Bernhard
TITLE OF INVENTION: Method for Generating an Immunogenic T
TITLE OF INVENTION: Cell Response Protective Against a Virus
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., P.O. Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/139,609
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/868,946
FILING DATE: 15-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/685,459
FILING DATE: 12-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/047,443
FILING DATE: 08-MAY-1987

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/725,087
FILING DATE: 19-APR-1985
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST1DUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-139-609-1

Query Match 10.5%; Score 99; DB 2; Length 369;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLLEDPV 177
DB 265 QPELAPEDPEDSALLLEDPV 283

RESULT 8
5470718-5
Patent No. 5470718
APPLICANT: O'CALLAGHAN, DENNIS J.
TITLE OF INVENTION: EQUINE HERPESVIRUS TYPE 1 GLYCOPROTEIN
D NUCLEIC ACIDS
NUMBER OF SEQUENCES: 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/954,417
FILING DATE: 30-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 561,553
FILING DATE: 01-AUG-1990
SEQ ID NO: 5
LENGTH: 383
5470718-5

Query Match 10.5%; Score 99; DB 6; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLLEDPV 177
DB 291 QPELAPEDPEDSALLLEDPV 309

RESULT 9
US-08-499-568-11
Sequence 11, Application US/08499568
Patent No. 5654174
GENERAL INFORMATION:
APPLICANT: Cohen, Gary H.
APPLICANT: Eisenberg, Roselyn J.
APPLICANT: Nicola, Anthony
TITLE OF INVENTION: Herpes Simplex Virus Glycoprotein D
TITLE OF INVENTION: Variants
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/499,568
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 5654174and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32813
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-499-568-11

Query Match 10.5%; Score 99; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLBDPV 177
Db 290 QPELAPEDPEDSALLBDPV 308

RESULT 10
US-08-793-958-11
Sequence 11, Application US/08793958
Patent No. 5814486
GENERAL INFORMATION:
APPLICANT: Cohen, Gary H.
APPLICANT: Bisenberg, Roselyn J.
TITLE OF INVENTION: Herpes Simplex Virus Glycoprotein D
TITLE OF INVENTION: Variants
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,958
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/499,568
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5814486and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32813
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-793-958-11

Query Match 10.5%; Score 99; DB 2; Length 388;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLBDPV 177
Db 290 QPELAPEDPEDSALLBDPV 308

RESULT 11
US-08-220-151-23
Sequence 23, Application US/08220151
Patent No. 5529780
GENERAL INFORMATION:
APPLICANT: Paolletti, Enzo
APPLICANT: Limbach, Keith J.
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtiss, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/220,151
FILING DATE: 30-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2540
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066 CURTMS
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-220-151-23

Query Match 10.5%; Score 99; DB 1; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLBDPV 177
Db 289 QPELAPEDPEDSALLBDPV 307

RESULT 12
US-08-413-118-23
Sequence 23, Application US/08413118
Patent No. 5688920
GENERAL INFORMATION:
APPLICANT: PAOLETTI, ENZO

APPLICANT: LIMBACH, KEITH J.
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC, AND GD AND USES THEREFOR
NUMBER OF SEQUENCES: 128
CORRESPONDENCE ADDRESS:
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
STREET: 530 FIFTH AVENUE, 25TH FLOOR
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/413,118
FILING DATE: 29-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,151
FILING DATE: 30-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: FROMMER, WILLIAM S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2670
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-413-118-23

Query Match 10.5%; Score 99; DB 1; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLEDPV 177
DB 289 QPELAPEDPEDSALLEDPV 307

RESULT 13
US-08-473-446-23
Sequence 23, Application US/08473446
Patent No. 6017542
GENERAL INFORMATION:
APPLICANT: PAOLETTI, ENZO
APPLICANT: LIMBACH, KEITH J.
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC, AND GD AND USES THEREFOR
NUMBER OF SEQUENCES: 128
CORRESPONDENCE ADDRESS:
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
STREET: 530 FIFTH AVENUE, 25TH FLOOR
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,446

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/413,118
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FROMMER, WILLIAM S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2670
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-473-446-23

Query Match 10.5%; Score 99; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLEDPV 177
DB 289 QPELAPEDPEDSALLEDPV 307

RESULT 14
5182195-10
Patent No. 5182195
APPLICANT: NAKAHAMA, KAZUO; KAISHO, YOSHIIKO; YOSHIMURA, KOJI
TITLE OF INVENTION: METHOD FOR INCREASING USING PROTEASE
DEFICIENT YEASTS
NUMBER OF SEQUENCES: 71
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/269,140
FILING DATE: 09-NOV-1988
SEQ ID NO: 10:
LENGTH: 393
5182195-10

Query Match 10.5%; Score 99; DB 6; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLEDPV 177
DB 289 QPELAPEDPEDSALLEDPV 307

RESULT 15
US-08-499-568-4
Sequence 4, Application US/08499568
Patent No. 5654174
GENERAL INFORMATION:
APPLICANT: Cohen, Gary H.
APPLICANT: Eisenberg, Roselyn J.
APPLICANT: Nicola, Anthony
TITLE OF INVENTION: Herpes Simplex Virus Glycoprotein D
TITLE OF INVENTION: Variants
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/499,568
 FILING DATE:
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5654174and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 32813
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 394 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-499-568-4

Query Match 10.5%; Score 99; DB 1; Length 394;
 Best Local Similarity 100.0%; Pred. No. 0.0022;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 159 QPELAPEDPEDSALIEDPV 177
 Db 290 QPELAPEDPEDSALIEDPV 308

Search completed: January 22, 2004, 11:38:43
 Job time : 34.2253 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 11:37:25 ; Search time 32.6212 Seconds
(without alignments)
521.804 Million cell updates/sec

Title: US-09-890-806-5

Perfect score: 947
Sequence: 1 MKFLVNVALVFMVVYISYIY.....GQDELAPEDEPEDSALLEDPV 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	740	78.1	199	1	QOBECS	HQLF2 protein prec
2	125	13.2	186	1	QOBECS	HQLF1 precursor -
3	102	10.8	70	1	MPHBI	melittin, major, p
4	99	10.5	393	1	VGBEDZ	glycoprotein D pre
5	99	10.5	394	1	VGBED1	glycoprotein D - h
6	99	10.5	394	1	VGBE17	glycoprotein D pre
7	99	10.5	394	1	A47627	glycoprotein D pre
8	96.5	10.2	1409	1	OFFPCP	copia polyprotein
9	91	9.6	393	1	VGBED2	glycoprotein D - h
10	91	9.6	393	2	VGBE33	glycoprotein D pre
11	91	9.6	393	2	E43674	US6 protein - huma
12	84.5	8.9	787	2	PC1232	copia polyprotein
13	84.5	8.9	1578	2	AD1512	peptidoglycan boun
14	84	8.9	159	2	F75021	ribosomal protein
15	84	8.9	254	2	AD0205	probable N-acetyl
16	81	8.6	410	2	S28354	hypothetical prote
17	80.5	8.5	415	2	T21532	hypothetical prote
18	80.5	8.5	1582	2	AC1153	adhesin homolog lm
19	80	8.4	703	1	ALBSX1	cyclomaltoedextrin
20	79	8.3	159	1	E71208	hypothetical prote
21	79	8.3	303	2	AI3215	hypothetical prote
22	78.5	8.3	404	2	S56535	conserved hypotet
23	78.5	8.3	404	2	E91287	hypothetical 43.7K
24	78.5	8.3	404	2	H86128	hypothetical prote
25	77.5	8.2	660	2	S71276	beta-fructofuranos
26	77.5	8.2	664	2	E86257	beta-fructosidase
27	77	8.1	272	2	T44528	phosphoglycolate p
28	76.5	8.1	248	1	A46542	granulin - Trichop
29	76	8.0	407	2	C82804	conserved hypotet

30	75.5	8.0	301	2	E72384	ABC transporter, A
31	75.5	8.0	316	2	A59021	aldehyde reductase
32	75.5	8.0	396	2	T04224	hypothetical prote
33	75.5	8.0	527	2	C70397	periplasmic cell d
34	75	7.9	394	2	AB1870	hypothetical prote
35	74.5	7.9	359	2	AI0140	moiyddenum transpo
36	74	7.8	261	2	A98950	hypothetical prote
37	74	7.8	261	2	E85798	hypothetical prote
38	74	7.8	351	2	C64948	probable membrane
39	74	7.8	361	2	S11237	polymerase - Berne
40	74	7.8	360	2	AI2854	conserved hypotet
41	74	7.8	360	2	F97631	hypothetical prote
42	74	7.8	473	1	T26280	11noleyl-CoA desa
43	73.5	7.8	557	2	T07909	4-coumarate-CoA li
44	73.5	7.8	757	2	B90572	lipoprotein (impor
45	73	7.7	463	2	D41853	hexose phosphate t

ALIGNMENTS

RESULT 1
QOBECS
HQLF2 protein precursor - human cytomegalovirus (strain AD169)
N/Alternate names: hypothetical protein US2
C/Species: human cytomegalovirus, human herpesvirus 5
C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 28-Jul-2000
C/Accession: E26078; S09916
R/Weston, K.; Barrell, B.G.
J. Mol. Biol. 192, 177-208, 1986
A/Title: Sequence of the short unique region, short repeats, and part of the long repeat:
A/Reference number: A92935; MUID:87169717; PMID:3031311
A/Accession: E26078
A/Molecule type: DNA
A/Residues: 1-199 <WES>
A/Cross-references: EMBL:X04650; NID:g59801; PIDN:CAB37096.1; PID:g4456177
A/Experimental source: Strain AD169
R/Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; J
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A/Reference number: S09749; MUID:90269039; PMID:2161319
A/Accession: S09916
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-199 <CHE>
A/Cross-references: EMBL:X17403; NID:g59591; PIDN:CAA35313.1; PID:g1780933
A/Experimental source: Strain AD169
A/Note: this sequence was submitted to the EMBL Data Library, December 1989
C/Genetics:
A/Gene: HQLF2
C/Superfamily: cytomegalovirus HQLF2 protein
C/Keywords: glycoprotein; transmembrane protein
F/1-25/Domain: signal sequence #status predicted <SIG>
F/26-199/Product: hypothetical protein US2 #status predicted <MAT>
F/161-186/Domain: transmembrane #status predicted <TMM>
F/68/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 78.1%; Score 740; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.1e-64;
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 RLPGDITKAGBDALRPWKSTAKHPWFQIEDNRCYIDNGKLPARGSIVGNMSRFVDPKAD 81
Db 20 RLPGDITKAGBDALRPWKSTAKHPWFQIEDNRCYIDNGKLPARGSIVGNMSRFVDPKAD 79
QY 82 YGIVENLYVHADDERVPGESIKMNVRLDVMPIFETLARLVLOGVIMRCVPELRV 141
Db 80 YGIVENLYVHADDERVPGESIKMNVRLDVMPIFETLARLVLOGVIMRCVPELRV 139
QY 142 DYTSSAYMMNMQYGMV 157
Db 140 DYTSSAYMMNMQYGMV 155

RESULT 4

VGEBDZ

glycoprotein D precursor - human herpesvirus 1 (strain Hzt)

C:Species: human herpesvirus 1

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Sep-1994

C:Accession: A90945; A03729

R:Laaky, L.A.; Dowbenko, D.J.

DNA 3, 23-29, 1984

A:Title: DNA sequence analysis of the type-common glycoprotein-D genes of herpes simplex

A:Reference number: A90945; PMID:84131549; PMID:6321120

A:Accession: A90945

A:Molecule type: DNA

A:Residues: 1-393 <LAS>

A:Cross-references: GB:K02372

C:Superfamily: herpesvirus glycoprotein D

C:Keywords: glycoprotein; transmembrane protein

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-393/Product: glycoprotein D #status predicted <GPD>

F:341-360/Domain: transmembrane #status predicted <TMN>

F:118,145,286/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.5%; Score 99; DB 1; Length 393;

Best Local Similarity 100.0%; Pred. No. 0.076;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 159 QPELAPEDPEDSALLDPV 177

DB 289 QPELAPEDPEDSALLDPV 307

RESULT 5

VGEBD1

glycoprotein D - human herpesvirus 1

C:Species: human herpesvirus 1

C:Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 16-Jul-1999

C:Accession: A94268; B90945; A03729

R:Watson, R.J.; Weis, J.H.; Salestrom, J.S.; Enquist, L.W.

Science 218, 381-384, 1982

A:Title: Herpes simplex virus type-1 glycoprotein D gene: nucleotide sequence and expres

A:Reference number: A94268; PMID:83016630; PMID:6289440

A:Accession: A94268

A:Molecule type: DNA

A:Residues: 1-394 <WAT>

A:Cross-references: GB:J02217; NID:g330100; PIDN:AAA45785.1; PID:g330101

A:Experimental source: strain Patton

A:Note: a strongly hydrophobic region of 25 amino acids between residues 340 and 364 is

R:Laaky, L.A.; Dowbenko, D.J.

DNA 3, 23-29, 1984

A:Title: DNA sequence analysis of the type-common glycoprotein-D genes of herpes simplex

A:Reference number: A90945; PMID:84131549; PMID:6321120

A:Accession: B90945

A:Molecule type: DNA

A:Residues: 1-3, 'A', 'S', '70', 'N', '72-83, 85-269, 'R', '271-282, 'P', '284-364, 'R', '366-394 <LAS>

A:Experimental source: strain Hzt

C:Superfamily: herpesvirus glycoprotein D

C:Keywords: glycoprotein; transmembrane protein

F:119,146,287/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.5%; Score 99; DB 1; Length 394;

Best Local Similarity 100.0%; Pred. No. 0.076;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 159 QPELAPEDPEDSALLDPV 177

DB 290 QPELAPEDPEDSALLDPV 308

RESULT 6

VGEB17

glycoprotein D precursor - human herpesvirus 1 (strain 17)

C:Species: human herpesvirus 1

A:Note: host Homo sapiens (man)

C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 16-Jun-2000

C:Accession: A03730

R:McGeoch, D.J.; Dolan, A.; Donald, S.; Rixon, F.J.

J. Mol. Biol. 181, 1-13, 1985

A:Title: Sequence determination and genetic content of the short unique region in the ger

A:Reference number: A00656; PMID:85160822; PMID:2984429

A:Accession: A03730

A:Molecule type: DNA

A:Residues: 1-394 <MCG>

A:Cross-references: GB:X14112; NID:g1944536; PIDN:CAA32283.1; PID:g59564

C:Superfamily: herpesvirus glycoprotein D

C:Keywords: glycoprotein; transmembrane protein

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-394/Product: glycoprotein D #status predicted <GPD>

F:342-358/Domain: transmembrane #status predicted <TMN>

F:119,146,287/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.5%; Score 99; DB 1; Length 394;

Best Local Similarity 100.0%; Pred. No. 0.076;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 159 QPELAPEDPEDSALLDPV 177

DB 290 QPELAPEDPEDSALLDPV 308

RESULT 7

A47627

glycoprotein D precursor - human herpesvirus 1 (strain ANG)

C:Species: human herpesvirus 1

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999

C:Accession: A47627

R:Rizumi, K.M.; Stevens, J.G.

J. Exp. Med. 172, 487-496, 1990

A:Title: Molecular and biological characterization of a herpes simplex virus type 1 (HSV-

A:Reference number: A47627; PMID:90324869; PMID:2165127

A:Accession: A47627

A:Molecule type: DNA

A:Residues: 1-394 <IZU>

A:Cross-references: GB:X54361; NID:g60414; PIDN:CAA38245.1; PID:g60415

C:Superfamily: herpesvirus glycoprotein D

C:Keywords: glycoprotein; transmembrane protein

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-394/Product: glycoprotein D #status predicted <GPD>

F:342-360/Domain: transmembrane #status predicted <TMN>

F:119,146,287/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.5%; Score 99; DB 1; Length 394;

Best Local Similarity 100.0%; Pred. No. 0.076;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 159 QPELAPEDPEDSALLDPV 177

DB 290 QPELAPEDPEDSALLDPV 308

RESULT 8

OFFFCP

copia polyprotein - fruit fly (Drosophila melanogaster) retrotransposon copia

N:Contains: copia protein, 31K; copia protein, 48K; proteinase

C:Species: Drosophila melanogaster

C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 16-Jul-1999

C:Accession: A03324; S03612; E14835

R:Mount, S.M.; Rubin, G.M.

Mol. Cell. Biol. 5, 1630-1638, 1985

A:Title: Complete nucleotide sequence of the Drosophila transposable element copia: homol

A:Reference number: A03324; PMID:85267679; PMID:2410772

A:Accession: A03324

A:Molecule type: DNA

A:Residues: 1-1409 <MOU>

A:Cross-references: GB:M1240; NID:g158615; PIDN:AAA74497.1; PID:g950318

R:Miller, K.; Rosenbaum, J.; Zbrzezna, V.; Pogo, A.O.

Nucleic Acids Res. 17, 2134, 1989
A>Title: The nucleotide sequence of Drosophila melanogaster copia-specific 2.1-kb mRNA.
A/Reference number: S03612; MUID:89183629; PMID:2538806
A/Accession: S03612
A/Molecule type: mRNA
A/Residues: 1-391,1375-1409 <MIL>
A/Cross-references: EMBL:X13719; NID:g7745; PIDN:CAA31997.1; PID:g7746
R.Yoshioaka, K.; Honma, H.; Zushi, M.; Kondo, S.; Togashi, S.; Miyake, T.; Shiba, T.
EMBO J. 9, 535-541, 1990
A>Title: Virus-like particle formation of Drosophila copia through autocatalytic process
A/Reference number: S14835; MUID:90151630; PMID:1689241
A/Accession: S14835
A/Molecule type: DNA
A/Residues: 1-391,1375-1409 <YOS>
A/Cross-references: EMBL:X54147; NID:g7749; PIDN:CAA38086.1; PID:g7750
C/Genetics:
A/Gene: FlyBase:copia
A/Cross-references: FlyBase:FBgn0000349
A/Mobile element: retrotransposon copia
C/Superfamily: retrovirus-related polyprotein
C/Keywords: polyprotein; proteinase
F:2-433/Product: copia protein, 48k #status predicted <MAT1>
F:2-270/Product: copia protein, 31k #status predicted <MAT2>
F:271-433/Product: proteinase #status predicted <MAT3>

Query Match 10.2%; Score 96.5; DB 1; Length 1409;
Best Local Similarity 23.8%; Pred. No. 0.63;
Matches 38; Conservative 27; Mismatches 48; Indels 47; Gaps 8;

QY 19 IYARLPDGTITKAGEDALRPWKS-----TAKHPWFQIED--NRCYIDNGKLFARGSIYGN 70
DB 1016 IYMLPQGISCSNDVCKLNKAIYGLKQARCFEVEFQALKECFVN-----SS 1065
QY 71 MSRFVFDPRADYGGVGENLY--HADVEFVPEGSLLKN-----VRNLDVMPFET 119
DB 1066 VDRCIY--LIDKGNINENITVLLVDDVVIATGDMTRMNNFKRYLMEKFRMTDLNEIKHF 1123
QY 120 LALRLVLOGDVIVLRCVPELRVDYTSAYM-----WNMQ 153
DB 1124 IGRIEMQEDKIYL-----SQSAVVKKILSKFNME 1153

RESULT 9
glycoprotein D - human herpesvirus 2
C/Species: human herpesvirus 2
C/Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 16-Jul-1999
C/Accession: A03731
R.Lasky, L.A.; Dowbenko, D.J.
DNA 3, 23-29, 1984
A>Title: DNA sequence analysis of the type-common glycoprotein-D genes of herpes simplex
A/Reference number: A90945; MUID:84131549; PMID:6321120
A/Accession: A03731
A/Molecule type: DNA
A/Residues: 1-393 <LAS>
A/Cross-references: GB:K02373; NID:g330270; PIDN:AAA45842.1; PID:g330271
C/Superfamily: herpesvirus glycoprotein D
C/Keywords: glycoprotein; transmembrane protein
F:119,146,287/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.6%; Score 91; DB 1; Length 393;
Best Local Similarity 94.4%; Pred. No. 0.46;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLLEDP 176
DB 290 QPELVPEDEPDSSALLLEDP 307

RESULT 10
glycoprotein D precursor - human herpesvirus 2 (strain 333)
C/Species: human herpesvirus 2

A/Note: host Homo sapiens (man)
C/Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
C/Accession: A03732
R.Watson, R.J.
Gene 26, 307-312, 1983
A>Title: DNA sequence of the herpes simplex virus type 2 glycoprotein D gene.
A/Reference number: A03732; MUID:84159516; PMID:6323270
A/Accession: A03732
A/Molecule type: DNA
A/Residues: 1-393 <WAT>
A/Cross-references: GB:K01408; NID:g330268; PIDN:AAA45841.1; PID:g330269
C/Superfamily: herpesvirus glycoprotein D
C/Keywords: glycoprotein; transmembrane protein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-393/Product: glycoprotein D #status predicted <GPD>
F:340-356/Domain: transmembrane #status predicted <TMN>
F:119,146,287/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.6%; Score 91; DB 1; Length 393;
Best Local Similarity 94.4%; Pred. No. 0.46;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLLEDP 176
DB 290 QPELVPEDEPDSSALLLEDP 307

RESULT 11
E43674
U56 protein - human herpesvirus 2 (strain HGS2)
C/Species: human herpesvirus 2
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C/Accession: E43674
R.McGeoch, D.J.; Moss, H.W.M.; McNab, D.; Frame, M.C.
J. Gen. Virol. 68, 19-38, 1987
A>Title: DNA sequence and genetic content of the HindIII 1 region in the short unique con
tiguous comparisons.
A/Reference number: A43674; MUID:87111457; PMID:3027242
A/Accession: E43674
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-393 <MCG>
A/Cross-references: EMBL:X04798
C/Superfamily: herpesvirus glycoprotein D

Query Match 9.6%; Score 91; DB 2; Length 393;
Best Local Similarity 94.4%; Pred. No. 0.46;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLLEDP 176
DB 290 QPELVPEDEPDSSALLLEDP 307

RESULT 12
PC1232
copia polyprotein - fruit fly (Drosophila simulans) retrotransposon copia (fragments)
C/Species: Drosophila simulans
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-Feb-1999
C/Accession: PC1232
R.Yoshioaka, K.; Kanda, H.; Takamatsu, N.; Togashi, S.; Kondo, S.; Miyake, T.; Sakaki, Y.;
Gene 120, 191-196, 1992
A>Title: Efficient amplification of Drosophila simulans copia directed by high-level rev
A/Reference number: PC1232; MUID:93013034; PMID:1383092
A/Accession: PC1232
A/Molecule type: DNA
A/Residues: 1-313;314-787 <YOS>
A/Cross-references: DDBJ:D10880
C/Genetics:
A/Gene: FlyBase:copia
A/Cross-references: FlyBase:FBgn0012867
A/Mobile element: retrotransposon copia
C/Superfamily: retrovirus-related polyprotein

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 11:37:25 ; Search time 58.5973 Seconds
(without alignments)
779.478 Million cell updates/sec

Title: US-09-890-806-5
Perfect score: 947
Sequence: 1 MKFLVNVALLVFMVVYISYIY.....GQPELAPEDPEDSALLDPV 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: SP archaea:*
- 2: SP bacteria:*
- 3: SP fungi:*
- 4: SP human:*
- 5: SP invertebrate:*
- 6: SP mammal:*
- 7: SP_mhc:*
- 8: SP organelle:*
- 9: SP phage:*
- 10: SP plant:*
- 11: SP rodent:*
- 12: SP virus:*
- 13: SP vertebrate:*
- 14: SP unclassified:*
- 15: SP rvirus:*
- 16: SP bacteriaph:*
- 17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	489	51.6	203	12 Q8QRV5	Q8qrv5 chimpanzee
2	136	14.4	186	12 Q918F3	Q918f3 human cytom
3	133	14.0	186	12 Q918F0	Q918f0 human cytom
4	132	13.9	186	12 Q910T7	Q910t7 human cytom
5	131	13.8	186	12 Q918E4	Q918e4 human cytom
6	130	13.7	186	12 Q918E6	Q918e6 human cytom
7	127	13.4	196	12 Q8JKE7	Q8jke7 cercopithec
8	124	13.1	186	12 Q918E2	Q918e2 human cytom
9	121	12.8	186	12 Q918E9	Q918e9 human cytom
10	118	12.5	186	12 Q910G4	Q910g4 human cytom
11	115	12.1	186	12 Q8UZK6	Q8uzk6 human cytom
12	112.5	11.9	149	12 Q918F4	Q918f4 human cytom
13	109.5	11.6	149	12 Q8QRV4	Q8qrv4 chimpanzee
14	108.5	11.5	149	12 Q910V7	Q910v7 human cytom
15	107.5	11.4	149	12 Q918F1	Q918f1 human cytom
16	106.5	11.2	149	12 Q918E7	Q918e7 human cytom

17	101.5	10.7	149	12 Q69186	Q69186 human cytom
18	100.5	10.6	149	12 Q918E3	Q918e3 human cytom
19	99	10.5	394	12 Q8QRB7	Q8qrb7 human herpe
20	99	10.5	394	12 Q991M3	Q991m3 human herpe
21	99	10.5	394	12 Q69081	Q69081 human herpe
22	99	10.5	394	12 Q05059	Q05059 human herpe
23	99	10.5	394	12 Q05060	Q05060 human herpe
24	99	10.5	394	12 Q8QRB6	Q8qrb6 human herpe
25	99	10.5	394	12 Q69082	Q69082 human herpe
26	97.5	10.3	149	12 Q910D0	Q910d0 human cytom
27	96.5	10.2	1017	5 Q8T391	Q8t391 drosophila
28	91	9.6	393	12 Q69467	Q69467 herpes simp
29	91	9.6	394	12 Q91NH6	Q91nh6 human herpe
30	85	9.0	376	16 Q98FC4	Q98fc4 rhizobium 1
31	84.5	8.9	1409	5 Q08461	Q08461 drosophila
32	84.5	8.9	1578	16 Q92E25	Q92e25 listeria in
33	84	8.9	159	17 Q9UY00	Q9uy00 pyrococcus
34	84	8.9	254	16 Q8ZFL6	Q8zfl6 yersinia pe
35	83	8.8	318	16 Q93J07	Q93j07 streptomyce
36	82	8.7	1938	2 P70983	P70983 bacillus sp
37	80.5	8.5	404	16 Q8CVG5	Q8cvg5 escherichia
38	80.5	8.5	415	5 Q19894	Q19894 caenorhabdi
39	80.5	8.5	1582	16 Q8Y9A5	Q8y9a5 listeria mo
40	80	8.4	250	16 Q8P8X2	Q8p8x2 xanthomonas
41	80	8.4	725	2 Q59239	Q59239 bacillus sp
42	79.5	8.4	692	2 Q30565	Q30565 bacillus br
43	79	8.3	159	17 Q59596	Q59596 pyrococcus
44	79	8.3	325	16 Q8UDM0	Q8udm0 agrobacteri
45	79	8.3	329	4 Q8N225	Q8n225 homo sapien

ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT;	203 AA.
ID Q8QRV5			
AC Q8QRV5;			
DT 01-JUN-2002 (TREMBlrel. 21, Created)			
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)			
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)			
DE US2.			
OS Chimpanzee cytomegalovirus.			
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC Betaherpesvirinae; Cytomegalovirus.			
OX NCBI_TaxID=188763;			
RN [1]			
RP SEQUENCE FROM N.A.			
RA Davison A.J., Akter P., Dolan A., Wright K.M., Addison C.,			
RA Alencador D.J., Hayward G.S., McGeoch D.J.;			
RT "The human cytomegalovirus genome revisited."			
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AF480884; AAM00783.1; -			
SQ SEQUENCE 203 AA; 23262 MW; CA131FD182BB8300 CRC64;			

Query Match 51.6%; Score 489; DB 12; Length 203;
Best Local Similarity 63.4%; Pred. NO. 6.2e-41;
Matches 85; Conservative 24; Mismatches 23; Indels 2; Gaps 1;

QY	24 PDGITRAGEDALRPWKSTAKHPQIEDNRCYIDNGKLFARGSIVGNNSRFVFPDKADYG 83
DB	28 PDPHTEL--TSYKPKWSTATRPWTITDENRCHIENGOMFGRGVSGLTTFVFDKADYG 85
QY	84 GVGENTLVHADVVEFVPEBSLKNNVRLNDVMPIFETLRLVQGDVIMLRVCPBLRVDY 143
DB	86 GVGENTLAVRAEDVEFIAGENLIRVQYANVLPILQRLVRLVINGDVVWLTCVPEIRVDY 145
QY	144 TSSAYVMNMQYGMV 157
DB	146 TSNAYVMNMQYGMV 159

RESULT 2

```
Q918F3
ID Q918F3 PRELIMINARY; PRT; 186 AA.
AC Q918F3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE US3 protein.
GN US3.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22A;
RA Scott G.M., Barrell B.G., Oram J., Rawlinson W.D.;
RT "Characterization of transcripts from the human cytomegalovirus genes
RL TRL7, UL20a, UL36, UL65, UL94, US3 and US34."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413648; AAL14474.1; -
SQ SEQUENCE 186 AA; 21588 MW; 3B9E8C0CB54EAB70 CRC64;

Query Match 14.4%; Score 136; DB 12; Length 186;
Best Local Similarity 26.5%; Pred. No. 1.2e-05;
Matches 44; Conservative 31; Mismatches 53; Indels 38; Gaps 9;

QY 10 VEMVVISYIYARLPDGTITKAGEDALRPWKSTAKHPFQIEDNRCYIDNGKLFARGSIGV 69
DB 4 VLMLAIIAFLRLADSVPR---PLNVVSEIKSAHFVEENQCFHMGMLYFKGRMSG 59
QY 70 NMSRFVDPKADYGVG-----ENLYV-----HADVEFVPGESLKMNV-----RNL 111
DB 60 NFT-----KKHFVNVGIVSQSYMDRLQVSGEQYHDE----RGAYFEMNIGGYPSHTV 109
QY 112 DVMPIFETLALRLVQLGD-VIMLRCPVELRVDTSSAYMMNQYGM 156
DB 110 DMVDI--TLSTR---WGDPKRYAACVPQVRMDYSSQTINWYLQSRM 150

RESULT 3
Q918F0 PRELIMINARY; PRT; 186 AA.
ID Q918F0;
AC Q918F0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE US3 protein.
GN US3.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=21A;
RA Scott G.M., Barrell B.G., Oram J., Rawlinson W.D.;
RT "Characterization of transcripts from the human cytomegalovirus genes
RL TRL7, UL20a, UL36, UL65, UL94, US3 and US34."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413650; AAL14480.1; -
SQ SEQUENCE 186 AA; 21558 MW; 21D1DC5ADF43DD1B CRC64;

Query Match 14.0%; Score 133; DB 12; Length 186;
Best Local Similarity 25.9%; Pred. No. 2.4e-05;
Matches 43; Conservative 32; Mismatches 53; Indels 38; Gaps 9;

QY 10 VEMVVISYIYARLPDGTITKAGEDALRPWKSTAKHPFQIEDNRCYIDNGKLFARGSIGV 69
DB 4 VLVLAIILAVLFLRLADSVPR---PLNVVSEIKSAHFVEENQCFHMGMLYFKGRMSG 59
QY 70 NMSRFVDPKADYGVG-----ENLYV-----HADVEFVPGESLKMNV-----RNL 111
DB 60 NFT-----KKHFVNVGIVSQSYMDRLQVSGEQYHDE----RGAYFEMNIGGYPSHTV 109
```

```
QY 112 DVMPIFETLALRLVQLGD-VIMLRCPVELRVDTSSAYMMNQYGM 156
DB 110 DMVDI--TLSTR---WGDPKRYAACVPQVRMDYSSQTINWYLQSRM 150

RESULT 4
Q910T7 PRELIMINARY; PRT; 186 AA.
ID Q910T7;
AC Q910T7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE US3 protein.
GN US3.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17A, 47A, and 48A;
RA Scott G.M., Barrell B.G., Oram J., Rawlinson W.D.;
RT "Characterization of transcripts from the human cytomegalovirus genes
RL TRL7, UL20a, UL36, UL65, UL94, US3 and US34."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413649; AAL14477.1; -
DR EMBL; AF413646; AAL14468.1; -
DR EMBL; AF413647; AAL14471.1; -
SQ SEQUENCE 186 AA; 21556 MW; 34D58C17BE5EA06B CRC64;

Query Match 13.9%; Score 132; DB 12; Length 186;
Best Local Similarity 25.9%; Pred. No. 3.1e-05;
Matches 43; Conservative 32; Mismatches 53; Indels 38; Gaps 9;

QY 10 VEMVVISYIYARLPDGTITKAGEDALRPWKSTAKHPFQIEDNRCYIDNGKLFARGSIGV 69
DB 4 VLVLAIILAVLFLRLADSVPR---PLNVVSEIKSAHFVEENQCFHMGMLYFKGRMSG 59
QY 70 NMSRFVDPKADYGVG-----ENLYV-----HADVEFVPGESLKMNV-----RNL 111
DB 60 NFT-----KKHFVNVGIVSQSYMDRLQVSGEQYHDE----RGAYFEMNIGGYPSHTV 109
QY 112 DVMPIFETLALRLVQLGD-VIMLRCPVELRVDTSSAYMMNQYGM 156
DB 110 DMVDI--TLSTR---WGDPKRYAACVPQVRMDYSSQTINWYLQSRM 150

RESULT 5
Q918E4 PRELIMINARY; PRT; 186 AA.
ID Q918E4;
AC Q918E4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE US3 protein.
GN US3.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=83A;
RA Scott G.M., Barrell B.G., Oram J., Rawlinson W.D.;
RT "Characterization of transcripts from the human cytomegalovirus genes
RL TRL7, UL20a, UL36, UL65, UL94, US3 and US34."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413663; AAL14519.1; -
SQ SEQUENCE 186 AA; 21604 MW; 26134793F853A47D CRC64;

Query Match 13.8%; Score 131; DB 12; Length 186;
Best Local Similarity 24.4%; Pred. No. 3.9e-05;
```

	Matches	41;	Conservative	33;	Mismatches	52;	Indels	42;	Gaps	9
Qy	10	VFMVVVSYIYARLPDGI	TRKAGEDL	RPMK----	STAKHPFQIEDRCYIDNGKL	FAR	64			
		:: :	:	:	:	:	:	:	:	:
Dd	4	VLVALIALAVLFRLADSV-----	PWPLDVVSEIRSAHFRVEENQCWFHGM	LTFK	54					
Qy	65	GSIVGNNSRFVPDPKADYG	GVGENL-----	YHADDFEFPGESL	KMNV-----	R	109			
		:: :	:	:	:	:	:	:	:	:
Dd	55	GRMSGNFTEKHf--VNVGIVS	QSVMDRLOVSGEQYHDE----	RGAYFEWNIGCHPVTHT	107					
Qy	110	NLDVMPIFETLALRLVLQGD-VIMLR	CVPELRVDYTSSAYMMNQYM	156						
		:: :	:	:	:	:	:	:	:	:
Dd	108	TVDMDVI--TLSTR--WGDPK	KTYAACVPOVRMDYSSQTINWYLQRSM	150						
 RESULT 6										
ID	Q918E6	PRELIMINARY; PRT; 186 AA.								
AC	Q918E6;									
DT	01-DEC-2001 (TrEMBLrel. 19,	Created)								
DT	01-DEC-2001 (TrEMBLrel. 19,	Last sequence update)								
DT	01-DEC-2001 (TrEMBLrel. 19,	Last annotation update)								
DE	US3 protein.									
GN	US3.									
OS	Human cytomegalovirus.									
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;									
CC	Betaherpesvirinae; Cytomegalovirus.									
OX	NCBI_TaxID=10359;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=9A;									
RA	Scott G.M., Barrell B.G., Oram J., Rawlinson W.D.;									
RT	"Characterization of transcripts from the human cytomegalovirus genes									
RT	TRL7, UL20a, UL36, UL65, UL94, US3 and US34."									
RL	Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.									
DR	EMBL; AF413657; AALL4501.1; -. .									
SQ	SEQUENCE 186 AA; 21679 MW; 7C0318DABA6E891D CRC64;									
 Query Match 13.7%; Score 130; DB 12; Length 186; Best Local Similarity 24.5%; Pred. No. 4.9e-05;										
	Matches	40;	Conservative	36;	Mismatches	55;	Indels	32;	Gaps	8
Qy	10	VFMVVVSYIYARLPDGI	TRKAGEDL	RPMKSTAKHPFQIEDRCYIDNGKL	FARGSIYG	69				
		:: :	:	:	:	:	:	:	:	:
Dd	4	VMMLAIALAVLFRLADSVPR	LDVTV---SEIRSAHFRVEENQCWFHGM	LTFKGRMG	59					
Qy	70	NMSRPFVEDPKADYG	GVGENL-----	YHADDFEFPGESL	KMNV-----	RNLDM	114			
		:: :	:	:	:	:	:	:	:	:
Dd	60	NFTKEHF--VNVGIVS	QSYMDRLOVSGEQYHDE----	RGAYFEWNIGCHPVTHT	112					
Qy	115	PIFETLALRLVLQGD-VIMLR	CVPELRVDYTSSAYMMNQYM	156						
		:: :	:	:	:	:	:	:	:	:
Dd	113	DI--TLSTR--WGDPK	KTYAACVPOVRMDYSSQTINWYLQRSM	150						
 RESULT 7										
ID	Q8JKE7	PRELIMINARY; PRT; 196 AA.								
AC	Q8JKE7;									
DT	01-OCT-2002 (TrEMBLrel. 22,	Created)								
DT	01-OCT-2002 (TrEMBLrel. 22,	Last sequence update)								
DT	01-MAR-2003 (TrEMBLrel. 23,	Last annotation update)								
DE	US2 protein.									
GN	US2.									
OS	Cercopithecine herpesvirus 8.									
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;									
CC	Betaherpesvirinae; Cytomegalovirus.									
OX	NCBI_TaxID=47929;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RA	MEDLINE=22174937; PubMed=12186931;									
RX	Chang W.L., Tarental A.F., Zhou S.S., Borowsky A.D., Barry P.A.;									
RT	"A recombinant thesus cytomegalovirus expressing enhanced green									

```

RT fluorescent protein retains the wild-type phenotype and pathogenicity
RT in fetal macaques."
RL J. Virol. 76:9493-9504(2002) .
DR EMBL; AF474179; AAM97606.1; -.
SQ SEQUENCE 196 AA; 23115 MW; A2195D7A67D809D CRC64;

Query Match          13.4%; Score 127; DB 12; Length 196;
Best local Similarity 26.1%; Pred. No. 0.0001;
Matches 29; Conservative 28; Mismatches 50; Indels 4; Gaps 4;

QY      42 AKHPWEIEDNRCYIDNGKLFARGSIVGNMSRFVDPKADYGVG-ENLYHADVEFVP 100
DB       |::: ||| ||| :||| ::||| ||| ::| |::| |::| |::|
        37 ARRALTVTQQTACYLEGGKLFMTGYITGIDSY-FVVVVYVRDKXIEDLOFFEKDMK-LS 94
         |::: ||| ::| |::| |::| |::| |::| |::| |::| |::|
QY      101 GESLKNNVRNLDMPIRETLALRLVL-OGDVILRCVPELRVDYTSSAYVM 150
DB       |::: ||| ::| |::| |::| |::| |::| |::| |::| |::|
        95 SDKLEFDLRYIEIDWTTTKVEMRFSLNNSDIIWTCERPHVKPDLTHNYLM 145

RESULT 8
QY18E2    PRELIMINARY; PRT; 186 AA.
ID QY18E2
AC QY18E2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE US3 protein.
GN US3.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=95A;
RA Scott G.M., Barrell B.G., Oram J., Rawlinson W.D.;
RT "Characterization of transcripts from the human cytomegalovirus genes
RT trL7, UL20a, UL36, UL65, UL94, US3 and US34.";
RL Submitted (Aug-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF413665; AAL14524.1; -.
SQ SEQUENCE 186 AA; 21602 MW; 3BE0504768183511 CRC64;

Query Match          13.1%; Score 124; DB 12; Length 186;
Best local Similarity 23.9%; Pred. No. 0.00019;
Matches 39; Conservative 36; Mismatches 56; Indels 32; Gaps 8;

QY      10 VFMVVYSIYIARLPDGITKAGEDALRPWKSTAGHPWFQIEDNRCYIDNGKLFGARSIVG 69
DB       |::: ||| ||| ::||| ::||| ::||| ::||| ::||| ::|||
        4 VLVALIAVLEFLRLADSVPRPLDVV---SEIRSAHFVEENQCWFHMGLYFKGRMSG 59
         |::: ||| ||| ::||| ::||| ::||| ::||| ::||| ::|||
QY      70 NMSRFVDPKADYGVGENL-----YHADVVEFVPGESLKMNV-----RNLDM 114
DB       |::: ||| ||| ::||| ::||| ::||| ::||| ::||| ::|||
        60 NFERHF--VNVGIVSKSYMRLQVSGEQYHDE---RGAYFEWNIGGHPTHTVDMV 112
         |::: ||| ||| ::||| ::||| ::||| ::||| ::||| ::|||
QY      115 PIKETLALRLVLQG-DVIWLRCVPELRVDYTSSAYMMNQYM 156
DB       |::: ||| ||| ::||| ::||| ::||| ::||| ::||| ::|||
        113 DI-TLSTR---WGDPKKYAACPQVRNDYSSQTINMYLQRSM 150

RESULT 9
QY18E9    PRELIMINARY; PRT; 186 AA.
ID QY18E9
AC QY18E9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE US3 protein.
GN US3.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
```

RP SEQUENCE FROM N.A.
RC STRAIN=27A;
RA Scott G.M., Barrell B.G., Oram J., Rawlinson W.D.;
RT "Characterization of transcripts from the human cytomegalovirus genes
TRL7, UL20a, UL36, UL65, UL94, US3 and US34."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413651; AAL14483.1; -
SQ SEQUENCE 186 AA; 21533 MW; CB68FE1E6BC76728 CRC64;
Query Match 12.8%; Score 121; DB 12; Length 186;
Best Local Similarity 24.2%; Pred. No. 0.00039;
Matches 39; Conservative 32; Mismatches 62; Indels 28; Gaps 7;
QY 10 VFMVYISYIYARLPDGTITAGEDALRPWKSTAKHPFQIEDNRCYIDNGKLFARGSIGV 69
DB 4 VLVLAIALAVLFLRLADSVPRPLDVV-----SEIRSAHFVEENQCFHMGMLHYKGRMSG 59
QY 70 NMSR-----FVEDPKADYGVGENLYVHADDFEVPGESLKNV-----RNLDPVPI 116
DB 60 NFTEKHFSVSGISQSYMDRLQVSGEQYHDE-----RGAYFEWNIGGHPVPTVDVMDI 114
QY 117 FETLALRLVLQGD-VIWLRCVPELRVDYTSAYMMNQYGM 156
DB 115 --TLSTR--WGDPKRYAACVPQVRMDYSSQTINWYLQ 150

RESULT 10
Q910G4 PRELIMINARY; PRT; 186 AA.
ID Q910G4;
AC Q910G4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE US3 protein.
GN US3.
OS Human cytomegalovirus (strain Towne), and
OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10363, 10359;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human cytomegalovirus; STRAIN=29A;
RA Scott G.M., Barrell B.G., Oram J., Rawlinson W.D.;
RT "Characterization of transcripts from the human cytomegalovirus genes
TRL7, UL20a, UL36, UL65, UL94, US3 and US34."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413652; AAL14486.1; -
DR EMBL; AY038933; AAK71882.1; -
SQ SEQUENCE 186 AA; 21515 MW; CB68FE1E6D6B0DE8 CRC64;

Query Match 12.5%; Score 118; DB 12; Length 186;
Best Local Similarity 24.1%; Pred. No. 0.00078;
Matches 38; Conservative 32; Mismatches 60; Indels 28; Gaps 7;
QY 10 VFMVYISYIYARLPDGTITAGEDALRPWKSTAKHPFQIEDNRCYIDNGKLFARGSIGV 69
DB 4 VLVLAIALAVLFLRLADSVPRPLDVV-----SEIRSAHFVEENQCFHMGMLHYKGRMSG 59
QY 70 NMSR-----FVEDPKADYGVGENLYVHADDFEVPGESLKNV-----RNLDPVPI 116
DB 60 NFTEKHFSVSGISQSYMDRLQVSGEQYHDE-----RGAYFEWNIGGHPVPTVDVMDI 114
QY 117 FETLALRLVLQGD-VIWLRCVPELRVDYTSAYMMNQ 153
DB 115 --TLSTR--WGDPKRYAACVPQVRMDYSSQTINWYLQ 147

RESULT 11
Q8UZK6 PRELIMINARY; PRT; 186 AA.
ID Q8UZK6;
AC Q8UZK6;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE US3 protein.
OS Human cytomegalovirus (strain Towne).
OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Towne;
RA Crew M.D.;
RT "Immune evasion genes from the Towne strain of human
cytomegalovirus."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY072774; AAL67142.1; -
SQ SEQUENCE 186 AA; 21335 MW; 86D0167FD66BAB84 CRC64;

Query Match 12.1%; Score 115; DB 12; Length 186;
Best Local Similarity 24.1%; Pred. No. 0.0015;
Matches 38; Conservative 32; Mismatches 60; Indels 28; Gaps 7;

QY 10 VFMVYISYIYARLPDGTITAGEDALRPWKSTAKHPFQIEDNRCYIDNGKLFARGSIGV 69
DB 4 VLVLAIALAVLFLRLADSVPRPLDVV-----SEIRSAHFVEENQCFHMGMLHYKGRMSG 59
QY 70 NMSR-----FVEDPKADYGVGENLYVHADDFEVPGESLKNV-----RNLDPVPI 116
DB 60 NFTEKHFSVSGISQSYMDRLQVSGEQYHDE-----RGAYFEWNIGGHPVPTVDVMDI 114
QY 117 FETLALRLVLQGD-VIWLRCVPELRVDYTSAYMMNQ 153
DB 115 --TLSTR--WGDPKRYAACVPQVRMDYSSQTINWYLQ 147

RESULT 12
Q918F4 PRELIMINARY; PRT; 149 AA.
ID Q918F4;
AC Q918F4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE US3 protein.
GN US3.
OS Human cytomegalovirus.
OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22A;
RA Scott G.M., Barrell B.G., Oram J., Rawlinson W.D.;
RT "Characterization of transcripts from the human cytomegalovirus genes
TRL7, UL20a, UL36, UL65, UL94, US3 and US34."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413648; AAL14475.1; -
SQ SEQUENCE 149 AA; 17217 MW; E822385AC2CE5420 CRC64;

Query Match 11.9%; Score 112.5; DB 12; Length 149;
Best Local Similarity 25.4%; Pred. No. 0.0021;
Matches 43; Conservative 32; Mismatches 49; Indels 45; Gaps 11;
QY 10 VFMVYISYIYARLPDGTITAGEDALRPWKSTAKHPFQIEDNRCYIDNGKLFARGSIGV 69
DB 4 VLMALIALAVLFLRLADSVPR-----PLNVVSEIKSAHFVEENQCFHMGMLYFKGRMSG 59
QY 70 NMSRFPVDPKADYGVG-----ENLVV-----HADDFEVPGESLKNV-----RNL 111

```

Db      60 NFT-----KKHFNVNIGVISQSYMDRLQVSGEQYHND-----RGAYFEWNIGGYPVSHTV
QY      112 DVMPIETLALRLVLQGD-VIMLRCPYELRNVYTSSAYM---NNMQYGM 156
      110 DMVDI--TLSTR---WGDPKKYACVPQVR-----TSWILKPTMARARYGL 149
Db

RESULT 13
O8QRV4
AC      O8ORV4      PRELIMINARY;      PRT;      187 AA.
DT      01-JUN-2002 (TREMBLrel. 21, Created)
DT      01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Immediate-early glycoprotein US3.
OS      Chimpanzee cytomegalovirus.
OC      Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC      Beroherpesvirinae; Cytomegalovirus.
OX      NCBI_TaxID=188763;
      [1]
RP      SEQUENCE FROM N.A.
RA      Davison A.J., Akter P., Dolan A., Wright K.M., Addison C.,
RA      Alencor D.J., Hayward G.S., McGeoch D.J.;
RT      "The human cytomegalovirus genome revisited.";
RL      Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF480884; AAM00784.1; -.
SQ      SEQUENCE 187 AA; 21636 MW; 24A8075A06C5E66C CRC64;

```

Query Match	11.6%;	Score 109.5;	DB 12;	Length 187;
Best Local Similarity	24.4%;	Pred. No. 0.0055;		
Matches	31;	Conservative	21;	Mismatches 44;
				Indels 31;
				Gaps 4

[illegible]

QY	150	WNMQYGM	156
Db	145	WYFORAM	151

RESULT 14
Q910V7
ID Q910V7 PRELIMINARY; PRT; 149 AA

DE US3111 protein.
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE US3111 protein.

05 Human cytomegalovirus.
0C viruses; dsDNA viruses, no RNA stage; Herpesviridae
0C Betaherpesvirinae; Cytomegalovirus.

ON - 100-100000
 RN [1]
 RP SEQUENCE FROM N.A.

RA	Scott G.M., Barrel
KC	SIRAIN=1/A, 4/A, 8

RT	"Characterization
----	-------------------

TRL7, UL20a, UL36,

Submitted (AUG-2000)

DR EMBL; AF413649; AA
DR EMBL; AF413646; AA

DR	EMBL; AF413646; AF
DR	EMBL; AF413647; AF

DR EMBL; AF413647; AA:
SQ SEQUENCE 149 AA;

SEQUENCE 149 AA;

Query Match

Best Local Similarity

Matches 42; Conserve

```

QY      10 VFMVVYIYIYAKLPDGIITKAGEDALRPMTKSTAKHPMFOIEDNRCYIDNGKLFARGSIVG 69
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      4 VLVLALIAVLFLRLADSVR----PLNVVSEIKSAHFRVEENQCMWFHMGMLYFKGRMSG 59
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      70 NMSRFVFDPKADGGVG-----ENLYV-----HADDVEFVPEGSILKNVY-----RNL 111
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      60 NFI-----KKHFVNIVGISQSYMDRLQVSGEYIHDE---RGAYFENNIGGYPVSHTV 109
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Qy      112 DVMPIETLALRLVLOGD-VIWRVCPELRVDYSSAYM---WNMOYGM 156
      ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db      110 DMVDI--TLSTR--WGDPKRYACVPQYR---TSWLYKPTWAARYGL 149

```

RESULT 15
Q918F1

AC	Q918F1;
DT	01-DEC-2001 (TREMBLrel. 19, Created)
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DE	US3iii protein.
GN	US3.
OS	Human cytomegalovirus.

OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]

RC STRAIN=21A;
RA Scott G.M., Barrell B.G., Oram J., Rawlinson W.D.,
RT "Characterization of transcripts from the human cy

RL Submitted (Aug-2001) to the EMBL/GenBank/DBD databases
DR EMBL; AF413650; AAL14481.1; -
SQ SEQUENCE 149 AA; 17201 MW; 0523494C741B3A79 CRC64;

Query Match	11.4%;	Score 107.5;	DB 12;	Length 149;
Best Local Similarity	24.9%;	Pred. No. 0.0065;		
Matches 42;	Conservative 33;	Mismatches 49;	Indels 45;	Gaps 11.

QY 10 VFMVVISYLARLPGLIKAGEDALRPWKSIAHAPWFQLEDNKKIIDLNGLLFAKRSIVG 69
| :: :: | | : : | | : :: | : : | : : : :
Db 4 VLVLALLAVLERLADSVPR----PLNVVSEIKSAHFREVEENQCFFHMGMALT YFKGRMSG 55

DQ NMSKFEVDFKADIGGVG-----ENLIY-----HADVEFEESESLNANV-----KNL ILL
| : || | : || :
Db 60 NPT-----KKHFNVGIVSQSYMDRLQVSSEGYHNDE---RCAYFEENNIGGHPRVTHTV 109

```
09      |::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
112 DMVFPIELMAKLVLDGV-VIWMKVCFEKLVDIISSAIM-----MNWVIGM    136
Db      ||||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
110 DNVDI--TLSTR---WGDPKKYAAACVPQVR----TSWYLKEPTVARYGL    149
```

Search completed: January 22, 2004, 11:41:34
Job time : 60.5973 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 11:37:25 ; Search time 17.4198 Seconds
(without alignments)
313.155 Million cell updates/sec

Title: US-09-890-806-3_COPY_28_143
Perfect score: 629
Sequence: 1 AGEDALRPWKSTAKHPWFQI.....LOGDVIMLRVCELRVDYTS 116

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	629	100.0	199	1	US02_HCMVA
2	97	15.4	186	1	US03_HCMVA
3	77.5	12.3	703	1	CDGT_BACCS2
4	74	11.8	261	1	ZNUB_ECOLI
5	72	11.4	1409	1	COP1_DROME
6	70.5	11.2	704	1	CDGT_BACCH
7	69.5	11.0	349	1	US30_HCMVA
8	69	11.0	292	1	FX16_HUMAN
9	67	10.7	381	1	GAL7_CRYNE
10	67	10.7	824	1	AD08_HUMAN
11	66.5	10.6	463	1	UHPT_SALTY
12	66.5	10.6	1151	1	ATC7_YEAST
13	66	10.5	212	1	KAD_STEMU
14	66	10.5	361	1	GLNA_PANAR
15	66	10.5	620	1	BCHD_CHLTE
16	65.5	10.4	286	1	ATPS_SCHPO
17	65.5	10.4	488	1	RNG_ECOLI
18	65.5	10.4	718	1	CDGT_BACCS
19	65.5	10.4	878	1	SVV_METUA
20	65	10.3	619	1	BCHD_CHLVI
21	64.5	10.3	192	1	VIF_HVIBS
22	64.5	10.3	272	1	GPHI_PSEAE
23	64.5	10.3	334	1	FX16_MOUSE
24	64.5	10.3	383	1	OP54_DROVI
25	64.5	10.3	840	1	EF2_ENTHI
26	64	10.2	229	1	Y997_HAEIN
27	64	10.2	313	1	GSHB_RHIL0
28	64	10.2	460	1	VP41_BPAPS
29	63.5	10.1	1059	1	CERU_RAT
30	63	10.0	286	1	SGBU_ECOLI
31	63	10.0	584	1	PME_BRANA
32	63	10.0	2252	1	POL1_GCMV
33	62.5	9.9	273	1	OTCC_MYCPN

34	62.5	9.9	378	1	OP54_DROME	P08255 drosophila
35	62.5	9.9	380	1	OP54_DROPS	P29404 drosophila
36	62.5	9.9	435	1	PEPC_LACLA	O9CEG3 lactococcus
37	62.5	9.9	458	1	ENGA_HELPL	O25505 helicobacte
38	62.5	9.9	462	1	ENGA_HELPU	O92109 helicobacte
39	62.5	9.9	517	1	VLI_HPV12	P36723 human papil
40	62.5	9.9	718	1	CDGT_BACCI	P30920 bacillus ci
41	62.5	9.9	997	1	Y414_MYCPN	P75183 mycoplasma
42	62.5	9.9	1062	1	CERU_MOUSE	O61147 mus musculu
43	62.5	9.9	1375	1	RPOB_VIBCH	O9KX30 vibrio chol
44	62	9.9	455	1	BLMH_CHICK	P87362 gallus gall
45	62	9.9	772	1	LPIG_DROME	P11997 drosophila

ALIGNMENTS

US02_HCMVA	STANDARD;	PRT;	199 AA.
ID	US02_HCMVA		
AC	P09712		
DT	01-MAR-1989 (Rel. 10, Created)		
DT	01-MAR-1989 (Rel. 10, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Hypothetical protein HQLF2.		
OS	US2.		
OS	Human cytomegalovirus (strain AD169).		
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Betaherpesvirinae; Cytomegalovirus.		
OX	NCBI_TaxID=10360;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=87169717; PubMed=3031311;		
RA	Weston K., Barrell B.G.;		
RT	"Sequence of the short unique region, short repeats, and part of the		
RT	long repeats of human cytomegalovirus.";		
RL	J. Mol. Biol. 192:177-208(1986)		
RN	[2]		
RP	COMPLETE GENOME.		
RX	MEDLINE=90269039; PubMed=2161319;		
RA	Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,		
RA	Horsnell T., Hutchison C.A. III, Kourarides T., Martignetti J.A.,		
RA	Prieddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;		
RT	"Analysis of the protein-coding content of the sequence of human		
RT	cytomegalovirus strain AD169.";		
RL	Curr. Top. Microbiol. Immunol. 154:125-169(1990).		
CC	-1- SIMILARITY: BELONGS TO THE US2 FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; X17403; CA35313.1; -		
DR	EMBL; X04650; CAB37096.1; -		
DR	PIR; E26078; QOBECS.		
DR	PDB; 1IM3; 27-JUN-01.		
KV	Hypothetical protein; 3D-structure.		
FT	CARBOHYD 68 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 172 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 188 N-LINKED (GLCNAC. . .) (POTENTIAL).		
SQ	SEQUENCE 199 AA; 23111 MW; 4DD2DF3D692393F3 CRC64;		
Query Match			
Best Local Similarity 100.0%; Score 629; DB 1; Length 199;			
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY	1 AGEDALRPWKSTAKHPWFQIEDNRCYIDNGTLFARGSVGNMSRFVDPKADYGVGENTL 60		
DB	28 AGEDALRPWKSTAKHPWFQIEDNRCYIDNGTLFARGSVGNMSRFVDPKADYGVGENTL 87		

QY 61 YHADVVEFVPGESLKMNVNLDVMPFETLALRLVLOGDVIMLRVCPBELRVDTYS 116
DB 88 YHADVVEFVPGESLKMNVNLDVMPFETLALRLVLOGDVIMLRVCPBELRVDTYS 143

RESULT 2

US03_HCMVA STANDARD; PRT; 186 AA.

ID US03_HCMVA STANDARD; PRT; 186 AA.

AC P09712; 01-MAR-1989 (Rel. 10, Last sequence update)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 01-FEB-1991 (Rel. 17, Last annotation update)

DE HQLF1 protein (E glycoprotein).

OS Human cytomegalovirus (strain AD169).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Cytomegalovirus.

NCBI_TaxID=10360;

SEQUENCE FROM N.A. MEDLINE=87169717; PubMed=3031311;

WESTON K., BARRELL B.G.; "Sequence of the short unique region, short repeats, and part of the

long repeats of human cytomegalovirus.";

J. Mol. Biol. 192:177-208 (1986).

SEQUENCE FROM N.A. MEDLINE=88128558; PubMed=2829427;

WESTON K.; "An enhancer element in the short unique region of human

cytomegalovirus regulates the production of a group of abundant

immediate early transcripts.";

Virology 162:406-416 (1988).

COMPLETE GENOME.

MEDLINE=90269039; PubMed=2161319; Bohm R., Brown C.M., Cerny R.,

Chen M.S., Bankier A.T., Beck S., Bohni R., Kozarides T., Martignetti J.A.,

Horrell T., Hutchison C.A. II, Kozarides T., Martignetti J.A.,

Predde E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;

"Analysis of the protein-coding content of the sequence of human

cytomegalovirus strain AD169.";

Curr. Top. Microbiol. Immunol. 154:125-169 (1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sib-sib.ch).

CC EMBL; X17403; CAA35314.1; -

DR EMBL; X04650; CAB37097.1; -

DR EMBL; M18921; AAA45958.1; -

DR PIR; F26078; QQBEC6.

KW Bary protein; Glycoprotein.

FT CARBOHYD 60 N-LINKED (GLCNAC..)(POTENTIAL).

FT SEQUENCE 186 AA; 21574 MW; 671753C1AA75920D CRC64;

QY 18 FQIEDNRCYIDNGKLFARSGIVGNSRFVDPKADYGGVGENL-----YHADVVE 68
DB 37 FRVVENQCWFHMGMLYFKGRSGNFTKGF---VNVGIVSGSYMDRLQVSGEGYHDE-- 91
QY 69 FVGEGLKMN-----RNLDMPIFETLALRLVLOGD-VIMLRVCPBELRVDTYS 116
DB 92 --RGAYFEWNIIGHPVTHTVDMVDI--TLSTR---WGDPKCYAACTVQVRMDYSS 139

RESULT 3

CDGT_BACS2 STANDARD; PRT; 703 AA.

ID CDGT_BACS2 STANDARD; PRT; 703 AA.

AC P1746; 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Cyclomaltoedextrin glucanotransferase precursor (EC 2.4.1.19)

DE (Cyclodextrin-glycosyltransferase) (CGTase).

GN CGT.

OS Bacillus sp. (strain 1-1).

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OC NCBI_TaxID=29334;

SEQUENCE FROM N.A., AND SEQUENCE OF 30-52.

Schmid G., Englbrecht A., Schmid D.;

"Cloning and nucleotide sequence of a cyclodextrin glycosyltransferase

gene from the alkalophilic Bacillus 1-1.";

(In) Huber O., Szejtli J. (eds.);

Proceedings of the fourth international symposium on cyclodextrins,

pp. 71-76, Kluwer Academic Publishers, Dordrecht and Boston (1988).

CC -1- CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation

of a 1,4-alpha-D-glucosidic bond.

CC -1- COFACTOR: BINDS TWO CALCIUM IONS.

CC -1- SUBUNIT: Monomer.

CC -1- MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE

IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND

IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER

ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN

ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE

MALTOOLIGOSACCHARIDE PRODUCED.

CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO

KNOWN AS THE ALPHA-AMYLASE FAMILY.

DR PIR; S26399; ALBSX1.

DR HSSP; P31797; 1CYG.

DR InterPro; IPR006589; Alp_amy1_cat_sub.

DR InterPro; IPR006048; Alpha_amy1_C-

DR InterPro; IPR006047; Alpha_amy1_cat.

DR InterPro; IPR002044; CBD 4.

DR InterPro; IPR006046; Glyco_hydro_13.

DR InterPro; IPR002909; IPT_TIG.

DR Pfam; PF00128; alpha-amyase; 1.

DR Pfam; PF02806; alpha-amyase_C; 1.

DR Pfam; PF00686; CBM 20; 1.

DR Pfam; PF01833; TIG_1.

DR PRINTS; PR00110; ALPHAAMYLASE.

DR PRODOM; PD001568; CBD 4; 1.

DR SMART; SM00642; Aamy; 1.

DR SMART; SM00632; Aamy_C; 1.

KW Transferrase; Glycosyltransferase; Calcium; Signal.

FT SIGNAL 1 29

FT CHAIN 30 703 CYCLOMALTOEDXTRIN GLUCANOTRANSFERASE.

FT DOMAIN 30 160 A1.

FT DOMAIN 161 224 B.

FT DOMAIN 225 428 A2.

FT DOMAIN 429 516 C.

FT DOMAIN 517 600 D.

FT DOMAIN 601 703 E.

FT DISULFID 68 75 BY SIMILARITY.

FT ACT SITE 251 251 BY SIMILARITY.

FT ACT SITE 279 279 BY SIMILARITY.

FT ACT SITE 350 350 BY SIMILARITY.

FT SEQUENCE 703 AA; 78663 MW; 4D973FB21D0D9B0A CRC64;

QY 12 TAKHPWFQIEDNRCYIDNGKLFARSGIVGNSRFVDPKADYGGVGENLYHADVVEFVP 71
DB 159 TENHSSPALETNPVVENGAITYDNGALLGN-----YSDNQNLPHHNGGIDFSS 207
QY 72 GE-SLKMNVRNLDVMPFETLALRLVLOGDVIMLRVCPBELRVDTYS 113

Db 208 YEDSIYRNLVLDADYDANNTVMDVYLKESIKFWLDKIGDIRVD 251

RESULT 4

ZNUB_ECOLI STANDARD; PRT; 261 AA.
AC P39832; P76286;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE High-affinity zinc uptake system membrane protein znub.
GN ZNUB OR B1859.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=562;
[1]

SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Berna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[2]

SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Sivasubram S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
[3]

SEQUENCE OF 1-92 FROM N.A.
RC STRAIN=K12 / EMG2;
RA Robison K., O'Keefe T., Church G.M.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
[4]

SEQUENCE OF 80-261 FROM N.A.
RX MEDLINE=88314937; PubMed=2842314;
RA Shinagawa H., Makino K., Anemura M., Kimura S., Iwaseki H., Nakata A.;
RT "Structure and regulation of the Escherichia coli ruv operon involved
in DNA repair and recombination.";
RL J. Bacteriol. 170:4322-4329(1988).
[5]

IDENTIFICATION.
RX MEDLINE=95075659; PubMed=7984428;
RA Borodovsky M., Rudd K.E., Koonin E.V.;
RT "Intrinsic and extrinsic approaches for detecting genes in a
bacterial genome.";
RL Nucleic Acids Res. 22:4756-4767(1994).
[6]

CHARACTERIZATION.
RX MEDLINE=98343803; PubMed=9680209;
RA Patzer S.I., Hantke K.;
RT "The ZnuABC high-affinity zinc uptake system and its regulator Zur in
Escherichia coli.";
RL Mol. Microbiol. 28:1199-1210(1998).
[7]

FUNCTION: INVOLVED IN THE HIGH-AFFINITY ZINC UPTAKE TRANSPORT
SYSTEM.
-1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(Probable).
-1- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE
PROTEINS. STRONG, TO H.INFLUENZAE ZNUB.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE000280; AAC74929.1; -.
DR EMBL; D90828; BAA15667.1; -.
DR EMBL; D90829; BAA15670.1; -.
DR EMBL; U38702; AAA81031.1; -.
DR EMBL; M21298; -; NOT_ANNOTATED_CDS.
DR PIR; C64948; C64948.
DR EcoGene; EG12368; znub.
DR InterPro; IPR001626; ABC_transp3.
DR Pfam; PF00950; ABC-3; 1.
KW Transport; Zinc transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 8 28 POTENTIAL.
FT TRANSMEM 54 74 POTENTIAL.
FT TRANSMEM 84 104 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 236 256 POTENTIAL.
SQ SEQUENCE 261 AA; 27728 MW; 78382B2EACE1490 CRC64;

Query Match 11.8%; Score 74; DB 1; Length 261;
Best Local Similarity 40.6%; Pred. No. 1.1;
Matches 13; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

Qy 82 LDVMPFETLRLVIGDVIWRCVPELRVD 113
Db 52 LDVNPYAVIAVTLTLGLGLVLEKRPQLAID 83

RESULT 5

COP1_DROME STANDARD; PRT; 1409 AA.
AC P04146; Q03728; Q24280; Q24555; Q24585; Q24586; Q24587;
DT 01-NOV-1986 (Rel. 03, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Copia protein [Contains: Copia VLP protein; Copia protease
(EC 3.4.23.-)].
GN COP1A.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
[1]

SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=85267679; PubMed=2410772;
RA Mount S.M., Rubin G.M.;
RT "Complete nucleotide sequence of the Drosophila transposable element
copia: homology between copia and retroviral proteins.";
RL Mol. Cell. Biol. 5:1630-1638(1985).
[2]

SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=85240569; PubMed=2409449;
RA Emori Y., Shiba T., Kanaya S., Inoue S., Yuki S., Saigo K.;
RT "The nucleotide sequences of copia and copia-related RNA in Drosophila
virus-like particles.";
RL Nature 315:773-776(1985).
[3]

SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=89183629; PubMed=2538806;
RA Miller K., Rosenbaum J., Zbrzezna V., Pogo A.O.;
RT "The nucleotide sequence of Drosophila melanogaster copia-specific
2.1-kb mRNA.";
RL Nucleic Acids Res. 17:2134-2134(1989).

```

RN [4]
RP SEQUENCE FROM N.A. (ISOFORM SHORT), AND MUTAGENESIS OF ASP-292.
RC TISSUE=Larva;
RX MEDLINE=90151630; PubMed=1689241;
RA Yoshioka K., Honma H., Zuehl M., Kondo S., Togashi S., Miyake T.,
RT Shiba T.;
RT "Virus-like particle formation of Drosophila copia through
RT autocatalytic processing.";
RL EMOB J. 9:535-541(1990).
CC -I- ALTERNATIVE PRODUCTS;
CC Event=Alternative splicing; Named isoforms=2;
CC Name=long;
CC IsoId=P04146-1; Sequence=Displayed;
CC Name=short;
CC IsoId=P04146-2; Sequence=VSP_005226;
CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A11.
CC -I- SIMILARITY: Contains 1 CCHC-type zinc finger.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X04456; CAA28054.2; -
DR EMBL; X02599; CAA26444.1; -
DR EMBL; X02599; CAA26445.1; -
DR EMBL; X02600; CAA26446.1; -
DR EMBL; X02600; CAA26447.1; -
DR EMBL; X13719; CAA31997.1; -
DR EMBL; X54147; CAA38086.1; -
DR PIR; A03324; OEFPCP.
DR MEROPS; A11.001; -.
DR FlyBase; FBgn0013437; copia\GIP.
DR InterPro; IPR001969; Aspprotease_site.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR001878; znf_CCHC.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; ZnF_C2HC; 1.
DR PROSITE; PS00141; ASP_PROTEASE; FALSE_NEG.
DR PROSITE; PS0158; ZF_CCHC; 1.
KW Transposable element; Hydrolase; Aspartyl protease; ATP-binding;
KW Polyprotein; Alternative splicing; Polymorphism; Zinc-finger.
FT CHAIN 1 270 COPIA VLP PROTEIN (POTENTIAL).
FT CHAIN 1 270 COPIA PROTEASE (POTENTIAL).
FT ZN_FING 230 247 CCHC-TYPE.
FT ACT_SITE 292 292 PROTEASE (BY SIMILARITY).
FT VARSPIC 392 1374 Missing (in isoform short).
FT VARIANT 1265 1288 /FTid=VSP_005226.
FT FTGTGLFKMFDELNLICWNTKRONs -> VOOGIVSKILLI
FT FTFTGIQDRITQ (IN VARIANT COPIA-RELATED).
FT MUTAGEN 1289 1409 MISSING (IN VARIANT COPIA-RELATED).
FT FT CONFLICT 191 191 D->A: LOSS OF ACTIVITY.
FT CONFLICT 300 300 S->N (IN REF. 2; CAA26447).
FT CONFLICT 300 300 I->V (IN REF. 2; CAA26447).
FT CONFLICT 866 866 Q->E (IN REF. 2; CAA26447).
SQ SEQUENCE 1409 AA; 162817 MM; BE89440763AA7691 CRC64;

```

ID	CDGT_BACOH	STANDARD;	PRT;	704 AA.
AC	P27036;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-NOV-1995 (Rel. 32, Last annotation update)			
DE	Cyclomalto-dextrin glucanotransferase precursor (EC 2.4.1.19)			
DE	(Cyclodextrin-glycosyltransferase) (CGTase).			
GN	CGT.			
OC	Bacillus ohbensis.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
RN	NCBI_TaxID=1481;			
RM	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92000599; PubMed=1368710;			
RA	Sin K.A., Nakamura A., Kobayashi K., Masaki H., Uozumi T.;			
RT	"Cloning and sequencing of a cyclodextrin glucanotransferase gene from Bacillus ohbensis and its expression in Escherichia coli.";			
RL	Appl. Microbiol. Biotechnol. 35:600-605(1991).			
CC	-1- CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation of a 1,4-alpha-D-glucosidic bond.			
CC	-1- COFACTOR: BINDS TWO CALCIUM IONS.			
CC	-1- SUBUNIT: Monomer.			
CC	-1- MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE MALTOLOGSACCHARIDE PRODUCED.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).			
CC	-----			
DR	EMBL; D90243; BAAL4289.2; .-			
DR	HSSP; P31797; 1CYG.			
DR	InterPro; IPR006589; Alp_amyl_cat_sub.			
DR	InterPro; IPR006048; Alpha_amyl_C.			
DR	InterPro; IPR006047; Alpha_amyl_cat.			
DR	InterPro; IPR002044; CBD_4.			
DR	InterPro; IPR006046; Glyco_hydro_13.			
DR	InterPro; IPR002909; IPT_TIG.			
DR	Pfam; PF00128; alpha-amylase; 1.			
DR	Pfam; PF02806; alpha-amylase_C; 1.			
DR	Pfam; PF00686; CBM_20; 1.			
DR	Pfam; PF01833; TIG; 1.			
DR	PRINTS; PR00110; ALPHAMYLAZE.			
DR	ProDom; PD001568; CBD_4; 1.			
DR	SMART; SM00642; Amyy_1.			
KW	Transferase; Glycosyltransferase; Calcium; Signal.			
FT	SIGNAL 1 29			
FT	CHAIN 30 704			
FT	DISULFID 68 75			
FT	ACT_SITE 251 251			
FT	ACT_SITE 279 279			
FT	ACT_SITE 350 350			
SEQ	SEQUENCE 704 AA; 78621 MM; 04FA14951D5ACEB CRC64;			
Query Match	11.2%; Score 70.5; DB 1; Length 704;			
Best Local Similarity	25.0%; Pred. No. 8.7;			
Matches 26; Conservative 21; Mismatches 44; Indels 13; Gaps 4;				

DB 159 TPNHSSPALETDSYAENGAVYNDGVLIGNYSN---DP-----NNLFPHNGGIDPFSS 207
QY 72 GE-SLKMNVRNLDVMPFETLALRLVLOGDVIWL-RCVPELRVD 113
DB 208 YEDSIYRNLVYLADYDLNNVTWMDQYLKESIKLMDKGIDGIRVD 251

RESULT 7

US30_HCMVA
ID US30_HCMVA STANDARD; PRT; 349 AA.
AC P09706;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Hypothetical protein HHRF5.
GN US30.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87169717; PubMed=3031311;
RA Weston K., Barrell B.G.;
RT "Sequence of the short unique region, short repeats, and part of the
long repeats of human cytomegalovirus.";
RL J. Mol. Biol. 192:177-208 (1986).
RN [2]
RP COMPLETE GENOME.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
cytomegalovirus strain AD169.";
RL Curr. Top. Microbiol. Immunol. 154:125-169 (1990).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; X17403; CAA35262.1; -.
DR EMBL; X04650; CAA28340.1; -.
DR PIR; E27216; QQBED5.
KW Hypothetical protein.
FT CARBOHYD 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 349 AA; 39114 MW; 613013B517DED750 CRC64;

Query Match 11.0%; Score 69.5; DB 1; Length 349;
Best Local Similarity 19.9%; Pred. No. 4.9;
Matches 29; Conservative 16; Mismatches 36; Indels 65; Gaps 5;

QY 24 RCYIDNGKLFARSGIVGNS-----RPVF-----DP 49
DB 81 QCRLRNATFSKEDIENSGPVVVELDYEDIDTGERQRLRPHLSGLGPTKENIRKDN 140
QY 50 KADY-GGVGENLYVHADYVF-----VPGESLKMNV-----79
DB 141 ESDVNGGIRWALYIQTDAKYGIRNGHLSIRLMYFGEKXTQQLDSDFSGRHRRSPSTPL 200
QY 80 -RNLVMPFETLALRLVLOGDVIWL 104
DB 201 GKNAEVPATRTSSTYSVLSAFVWVI 226

RESULT 8

FX16_HUMAN
ID FX16_HUMAN STANDARD; PRT; 292 AA.
AC Q81X29;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE F-box only protein 16.
GN FBXO16 OR FBX16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mao Y., Xie Y., Jiang M.;
RT "CDNA cloning of a novel human F-box protein.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 F-box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; AF453435; AAN76812.1; -.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR PROSITE; PS0181; FBOX; 1.
KW Ubl conjugation.
FT DOMAIN 86 F-BOX.
SQ SEQUENCE 292 AA; 34588 MW; 0A7030EFE682EDFD CRC64;

QY 1 AGEDALRPWKSTAKHPWFQIEDNRCYIDN 29
DB 211 SGEKALPWRSSDKHPDTIIRFN--YLDN 237
Query Match 11.0%; Score 69; DB 1; Length 292;
Best Local Similarity 51.7%; Pred. No. 4.5;
Matches 15; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

RESULT 9
GAL7_CRYNE
ID GAL7_CRYNE STANDARD; PRT; 381 AA.
AC P40908;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Galactose-1-phosphate uridylyltransferase (EC 2.7.7.12) (Gal-1-P
uridylyltransferase) (UDP-glucose--hexose-1-phosphate
uridylyltransferase).
GN GAL7.
OS Cryptococcus neoformans (Filobasidiella neoformans).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=5507;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96020643; PubMed=8577246;
RA Wickes B.L., Edman J.C.;
RT "The Cryptococcus neoformans GAL7 gene and its use as an inducible
promoter.";
RL Mol. Microbiol. 16:1099-1109 (1995).
CC -1- CATALYTIC ACTIVITY: UDP-glucose + alpha-D-galactose 1-phosphate =
alpha-D-glucose 1-phosphate + UDP-galactose.

```
CC -1 FUNCTION: POSSIBLE INVOLVEMENT IN EXTRAVASATION OF LEUCOCYTES.
CC -1 COFACTOR: Binds 1 zinc ion per subunit (Probable).
CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
CC -1 TISSUE SPECIFICITY: EXPRESSED ON NEUTROPHILS AND MONOCYTES.
CC -1 SIMILARITY: Belongs to peptidase family M12B.
CC -1 SIMILARITY: Contains 1 disintegrin domain.
CC -1 DATABASE: NAME=PROW; NOTE=CD guide Cd156 entry;
CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd156.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D26579; BAA05626.1; -.
DR HSSP, P18619; 1FVL.
DR MEROPS, M12.208; -.
DR Genew; HGNC:215; ADAM8.
DR MIM; 602267; -.
DR GO; GO:0005888; C:proteoglycan integral to plasma membrane; TAS.
DR InterPro; IPRO06586; ADAM_cysteine.
DR InterPro; IPRO01762; Disintegrin..
DR InterPro; IPRO06209; EGF like.
DR InterPro; IPRO06210; IEGF.
DR InterPro; IPRO02870; pep_M12B_propep.
DR InterPro; IPRO01590; Repolysin.
DR InterPro; IPRO06025; Zn_MTpeptidase.
DR Pfam; PF00200; disintegrin; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR PRINTS; PR00289; DISINTEGRIN.
DR Prodom; PD000664; Disintegrin; 1.
DR SMART; SMO00608; ACR; 1.
DR SMART; SMO0050; DISIN; 1.
DR SMART; SMO0181; EGF; 1.
DR PROSITE; PS50215; ADAM_MEROP; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloproteinase; zinc; Signal; Glycoprotein;
KW Transmembrane; Antigen.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 824 ADAM 8.
FT DOMAIN 17 655 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 656 676 POTENTIAL.
FT DOMAIN 677 824 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 200 400 METALLOPROTEASE.
FT DOMAIN 408 494 DISINTEGRIN-LIKE.
FT METAL 334 334 ZINC (CATALYTIC) (PROBABLE).
FT ACT_SITE 335 335 BY SIMILARITY.
FT METAL 338 338 ZINC (CATALYTIC) (PROBABLE).
FT METAL 344 344 ZINC (CATALYTIC) (PROBABLE).
FT DISULFID 310 395 BY SIMILARITY.
FT CARBOHYD 67 67 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 91 91 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 436 436 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 824 AA; 88673 MW; 5DF8E05F30DF479E CRC64;

Query Match 10.7%; Score 67; DB 1; Length 824;
Best Local Similarity 26.4%; Pred. No. 25;
Matches 29; Conservative 18; Mismatches 35; Indels 28; Gaps 6;
```

Db 203 ELYVVVDNAEFQMLGSEAAVRHRYLEVNVHDKLYQKINFRVVLGLEIW 252

RESULT 11

UHPT_SALTY STANDARD; PRT; 463 AA.

ID UHPT_SALTY STANDARD; PRT; 463 AA.

AC P27670;

DT 01-AUG-1992 (Rel. 23, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hexose phosphate transport protein.

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OX NCBI_TaxID=602;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=LT2;

RX MEDLINE=92234930; PubMed=1569007;

RA Island M.D., Wei B.-Y., Kadner R.J.;

RT "Structure and function of the uhp genes for the sugar phosphate transport system in Escherichia coli and Salmonella typhimurium.";

RL J. Bacteriol. 174:2754-2762(1992).

[2]

RN SEQUENCE FROM N.A.

RP STRAIN=LT2 / SGSC1412 / ATCC 700720;

RX MEDLINE=21534948; PubMed=11677609;

RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;

RT "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.";

RL Nature 413:852-856(2001).

CC -1- FUNCTION: TRANSPORT PROTEIN FOR SUGAR PHOSPHATE UPTAKE.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.

CC -1- INDUCTION: EXTERNAL GLUCOSE-6-PHOSPHATE INDUCES THE EXPRESSION OF THE UHP-REGION.

CC -1- SIMILARITY: BELONGS TO THE SLC37A FAMILY OF TRANSPORTERS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M89480; AAA27246.1; -.

DR EMBL; AE008876; AAL22645.1; -.

DR PIR; D41853; D41853.

DR StyGene; SG10409; uhpT.

DR InterPro; IPR000849; G1PT_transporter.

DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sub_transporter.

DR Pfam; PF00083; sugar_tr; 1.

DR TIGRFAMs; TIGR00881; 2A0104; 1.

DR PROSITE; PS00942; GLPT; 1.

KW Transport; Sugar transport; Transmembrane; Inner membrane;

KW Complete proteome.

FT DOMAIN 1 26 CYTOPLASMIC (PROBABLE).

FT TRANSMEM 27 45 PROBABLE.

FT DOMAIN 46 66 PERIPLASMIC (PROBABLE).

FT TRANSMEM 67 89 PROBABLE.

FT DOMAIN 90 96 CYTOPLASMIC (PROBABLE).

FT TRANSMEM 97 116 PROBABLE.

FT DOMAIN 117 120 PERIPLASMIC (PROBABLE).

FT TRANSMEM 121 140 PROBABLE.

FT DOMAIN 141 159 CYTOPLASMIC (PROBABLE).

FT TRANSMEM 160 187 PROBABLE.

FT DOMAIN 188 190 PERIPLASMIC (PROBABLE).

FT TRANSMEM 191 210 PROBABLE.

FT DOMAIN 211 249 CYTOPLASMIC (PROBABLE).

FT TRANSMEM 250 274 PROBABLE.

FT DOMAIN 275 294 PERIPLASMIC (PROBABLE).

FT TRANSMEM 295 318 PROBABLE.

FT DOMAIN 319 327 CYTOPLASMIC (PROBABLE).

FT TRANSMEM 328 344 PROBABLE.

FT DOMAIN 345 350 PERIPLASMIC (PROBABLE).

FT TRANSMEM 351 379 PROBABLE.

FT DOMAIN 380 391 CYTOPLASMIC (PROBABLE).

FT TRANSMEM 392 411 PROBABLE.

FT DOMAIN 412 425 PERIPLASMIC (PROBABLE).

FT TRANSMEM 426 446 PROBABLE.

FT DOMAIN 447 463 CYTOPLASMIC (PROBABLE).

FT CONFLICT 16 16 P -> A (IN REF. 1).

FT CONFLICT 385 386 GA -> AL (IN REF. 1).

FT CONFLICT 413 413 G -> A (IN REF. 1).

SO SEQUENCE 463 AA; 50708 MW; C13398B21CEA92DA CRC64;

Query Match 10.6%; Score 66.5; DB 1; Length 463;

Best Local Similarity 31.1%; Pred. No. 14;

Matches 23; Conservative 7; Mismatches 27; Indels 17; Gaps 4;

QY 38 IVGNNS-RFVDPKADYG-GVGENLY---VHADVDFVPGSLKMNVRNLDVMPIFETLA 92

Db 204 IVGFGLRFSGSDSPESYGLGKABELFGEEISEDXETENEMTKWQI-----F 251

QY 93 LRLVIGDVIWLRG 106

Db 252 VEYVKNKVIWLLC 265

RESULT 12

ATC7_YEAST STANDARD; PRT; 1151 AA.

ID ATC7_YEAST STANDARD; PRT; 1151 AA.

AC P40527;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Potential phospholipid-transporting ATPase 4 (EC 3.6.3.1).

GN NEO1 OR YIL048W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;

RX PubMed=9169870;

RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D., Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N., Harris D.B., Horsnell T., Hunt S., Jagsels K., Jones M., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N., Skelton J., Smith V., Walsh S., Whitehead S., Barrall B.G.;

RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";

RL Nature 387:84-87(1997).

CC -1- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE TRANSPORT OF PHOSPHOLIPIDS (POTENTIAL).

CC LEADS TO NEOMYCIN-RESISTANCE WHEN OVEREXPRESSED.

CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: Belongs to the cation transport ATPases family (P-type ATPases). Subfamily IV.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; Z38060; CAA86174.1; -.

```

DR PTR; S48431; S48431.
DR SGD; S0001310; NEOL.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR006539; Flippase.
DR InterPro; IPR005834; Hydrolase.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR TIGRFAMs; TIGR01652; ATPase-Plipid; 1.
DR TIGRFAMs; TIGR01494; ATPase_P-type; 7.
DR PROSITE; PS00154; ATPase_E1-E2; 1.
DR Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
FT TRANSMEM 368 388 POTENTIAL.
FT TRANSMEM 417 437 POTENTIAL.
FT TRANSMEM 439 459 POTENTIAL.
FT TRANSMEM 948 968 POTENTIAL.
FT TRANSMEM 971 991 POTENTIAL.
FT TRANSMEM 1021 1041 POTENTIAL.
FT TRANSMEM 1053 1073 POTENTIAL.
FT TRANSMEM 1079 1099 POTENTIAL.
FT TRANSMEM 1110 1130 POTENTIAL.
FT MOD_RES 503 503. PHOSPHORYLATION (PROBABLE).
SQ SEQUENCE 1151 AA; 130217 MW; DC7225CC9577DBE6 CRC64;

Query Match 10.6%; Score 66.5; DB 1; Length 1151;
Best Local Similarity 24.1%; Pred. No. 41;
Matches 26; Conservative 24; Mismatches 47; Indels 11; Gaps 5;

Oy LRPWKSTAKHPWFQIEDNRCYINGKLFARGSIGVNSRFPVDPKADYGVGENLYVHAD 65
Db IKIMWLTGCD---KVETARCVSISAKLISRGQYVHTITK-VTRPEGAFNQI-EYLKINRN 821
Oy DVEFVPEESLKNVNRNLDMPIFETLALRLVQGVYIWLRCVPELAVD 113
Db ACLLIDESLGMFLKHYE-QEFPDY---VWHLFVYIACRCTPQQRAD 864

RESULT 13
KAD_STRMU
ID_KAD_STRMU STANDARD; PRT; 212 AA.
AC Q8D633;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adenylate kinase (EC 2.7.4.3) (ATP-AMP transphosphorylase).
GN ADK OR SMU.2005.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1309;
OX [1]
RN
RP SEQUENCE FROM N.A.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; Pubmed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Prioleaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -!- FUNCTION: This small ubiquitous enzyme is essential for
CC maintenance and cell growth.
CC -!- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the adenylate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE015024; AANS59609.1; -.
DR HAMAP; MF_00235; -, 1.
DR Pfam; PF00406; ADK; 1.
DR ProDom; PD000657; Adenylate_kin; 1.
DR TIGRFAMs; TIGR01351; adk; 1.
DR PROSITE; PS00113; ADENYLATE KINASE; 1.
DR Transferrase; kinase; ATP-binding; Complete proteome.
KW NP BIND 7 ATP (BY SIMILARITY).
FT SEQUENCE 212 AA; 23638 MW; 5AF9C1AD9B5D5520 CRC64;
SQ
Query Match 10.5%; Score 66; DB 1; Length 212;
Best Local Similarity 23.7%; Pred. No. 6.5;
Matches 22; Conservative 18; Mismatches 27; Indels 26; Gaps 5;
Qy 22 DNRQYID-NGKLFGSGIVGNMSRFVDPKADYGGVGENLYVHAD----- 66
Db 117 DPACILVERLSGRITNRRKT-GETTYHKVNPADYN-EDDIYQREDDKPEYVKRRLDVNI 172
Qy 67 -----VEFVPGESLKMNVN-NDVMPFETLA 92
Db 173 AQGHPRIEYRNKGLVYDIEGNDINLVETIA 205
RESULT 14
GLNA_PANAR STANDARD; PRT; 361 AA.
ID GLNA_PANAR
AC Q04831;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase).
OS Panulirus argus (Spiny lobster).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Decapoda; Decapoda; Pleocyemata; Palinura;
OC Palinuroidea; Palinuridae; Panulirus.
OX NCBI_TaxID=6737;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory organ;
RX MEDLINE=93314973; PubMed=8100791;
RA Trapido-Rosenthal H.G., Linser P.J., Greenberg R.M., Gleason R.A.,
RA Carr W.B.;
RT "cDNA clones from the olfactory organ of the spiny lobster encode a
RT protein related to eukaryotic glutamine synthetase.";
RL Gene 129:275-278(1993).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -1- SUBUNIT: Homooctamer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M96798; AAA02583.1; -.
CC PIR; JN0716; JN0716.
CC HSSP; P06201; ILGR.
CC InterPro; IPR001691; GLN_synth.
CC InterPro; IPR001637; GLN_adenyltn.
CC Pfam; PF00120; gln-synt; 1.
CC Pfam; PF03951; gln-synt_N; 1.
CC ProDom; PD001057; gln_synt_C; 1.
CC PROSITE; PS00180; GLNA_1; 1.
CC PROSITE; PS00181; GLNA_ATP; 1.

```

```

KW      Ligase.          '
SQ      SEQUENCE        361 AA;  40768 MM;   3D8C3C507676099C CRC64;

Query Match           10.5%; Score 66; DB 1; Length 361;
Best Local Similarity 23.2%; Pred. No. 12;
Matches    29; Conservative    14; Mismatches    44; Indels    38; Gaps    6

QY      4 DALRPMK-----STAKHPMFQIEDNRCYIDNGKL--FARGSIGVGNMSRFYDPKADY 53
         | : ||| : ||||| : | : | : | : | : | : | : | : | : | : | : | : |
DB      107 DTNQRMKCMENVMTRAADQHPMFGMEQEYTLTDIDKHPLGWPKNGVPQGQPY-----Y 159

QY      54 GGVGENLYVHADDFEF-----VPGESLKNMVRNLDPWPI---FETLARLVLG 99
         ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      160 CGVGANRVYGRDVVEAHYRACLCAGINISGENAK-----VMPAQWFQVGPCEGITWG 212

QY      100 DVIVIL 104
         | : | :
DB      213 DDLMV 217

```

RESULT 15

```

ID   _BCHD_CHLITE          STANDARD;          PRT;          620 AA.
AC   Q935W0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Magnesium-chelatase 67 kDa subunit (Mg-protoporphyrin IX chelatase)
DE   (Mg-chelatase subunit D).
GN   BCHD OR CT1296.
OS   Bacterium tepidum.
OC   Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC   Chlorobium.
OX   NCBI_TaxID=1097;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20433268; PubMed=10976061;
RT   Xiong J., Fischer W.M., Inoue K., Nakahara M., Bauer C.E.;
RL   "Molecular evidence for the early evolution of photosynthesis.";
    Science 289:1724-1730(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   STRAIN=TLS / ATCC 49652 / DSM 12025;
RT   MEDLINE=22103685; PubMed=12093901;
RA   Eisen J.A., Neilson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA   Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA   Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA   Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA   Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA   Vanathavan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA   Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT   "The complete genome sequence of Chlorobium tepidum TLS, a
RT   photosynthetic, anaerobic, green-sulfur bacterium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC   -1- FUNCTION: Involved in bacteriochlorophyll biosynthesis; introduces
CC   a magnesium ion into protoporphyrin IX to yield Mg-protoporphyrin
CC   IX.
CC   -1- PATHWAY: Bacteriochlorophyll biosynthesis.
CC   -1- SIMILARITY: BELONGS TO THE MG-CHELATASE SUBUNITS D/I FAMILY.
CC   -1- SIMILARITY: Contains 1 WMPA domain.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   at the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AY005135; AAG12406.1; -.
DR   EMBL; AE012889; AAM72526.1; -.
DR   TIGR; CT1296; -.
DR   InterPro; IPR000523; Mg_chelatase_chII.

```

DR	InterPro:IPR002035; VWF_A.
DR	Pfam; PF01078; Mg_chelataase; 1.
DR	Pfam; PF00092; vwa; 1.
DR	PROSITE; PS0234; VWF_A; 1.
KW	Photosynthesis; Bacteriochlorophyll biosynthesis; Complete proteome.
FT	DOMAIN 277 329 GLU/PRO-RICH.
FT	DOMAIN 432 620 VWF_A.
SQ	SEQUENCE 620 AA; 67015 MW; 05BE07DD3D96F917 CRC64;
Query Match	10.5%; Score 66; DB 1; Length 620;
Best Local Similarity	29.2%; Pred. No. 23;
Matches 31; Conservative 10; Mismatches 37; Indels 28; Gaps 5,	
OY	3 EDALRPWKSTA-----KHPWFOIEDNRCTIDNGSKLF-----ARGSTVGNSHVFVDFPKAD 52
DB	401 EQAKRKIKSTALIIIGKD--DIKIKRFPRDKSGTLFIEMVDASGSMAIIRMR---QAQGA 454
OY	53 YGVGGENLYIHADDVEFY-----PGESLKMNVRLNDVMP 86
DB	455 VASLLONAYTRHDQVSLSIFRGKQAQVLPPSQSDVRAKELDLVP 500

Search completed: January 22, 2004, 11:39:40
Job time : 19.4198 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 11:37:25 ; Search time 64.6382 Seconds
(without alignments)
434.644 Million cell updates/sec

Title: US-09-890-806-5
Perfect score: 947
Sequence: 1 MKFLVNVALLVFMVVVYSIVY.....GQPELAPEDPEDSALLIEDPV 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	947	100.0	177	21	AAV97249
2	740	78.1	199	21	AAV97248
3	117.5	12.4	451	23	ABG31298
4	110	11.6	498	23	ABG31300
5	109.5	11.6	1189	19	AAW47008
6	106.5	11.2	354	21	AAAB10640
7	106.5	11.2	354	21	AAAB10641
8	104	11.0	386	22	AAE05128
9	104	11.0	496	23	ABG31502

10	102	10.8	21	22	AA63690
11	102	10.8	21	22	AA672437
12	102	10.8	21	23	ABG31292
13	102	10.8	21	23	ABG10034
14	102	10.8	21	23	AAU76918
15	102	10.8	21	23	AAW52328
16	102	10.8	21	24	ABU08142
17	102	10.8	22	21	AAV70301
18	102	10.8	48	23	ABH79912
19	102	10.8	49	21	AAH36293
20	102	10.8	129	24	ABG74777
21	102	10.8	350	21	AAV70278
22	99	10.5	124	23	AAW49703
23	99	10.5	308	18	AAW11322
24	99	10.5	369	11	AAW06489
25	99	10.5	369	20	AAW81601
26	99	10.5	388	18	AAW11323
27	99	10.5	393	10	AAW90887
28	99	10.5	394	5	AAW40083
29	99	10.5	394	6	AAW50582
30	99	10.5	509	23	AAW49704
31	98	10.3	21	23	AAW52506
32	97	10.2	20	23	ABG31288
33	97	10.2	21	17	AAW99533
34	97	10.2	21	20	AAW87499
35	91	9.6	368	16	AAW12068
36	91	9.6	393	5	AAW40084
37	91	9.6	393	10	AAW90748
38	91	9.6	393	18	AAW11324
39	91	9.6	393	19	AAW72163
40	91	9.6	394	6	AAW50581
41	91	9.6	400	19	AAW2143
42	91	9.6	400	19	AAW72013
43	87.5	9.2	26	19	AAW64613
44	87.5	9.2	26	19	AAW61590
45	87.5	9.2	26	20	AAW85766

ALIGNMENTS

RESULT 1
AAV97249 standard; Protein; 177 AA.

AAV97249;

04-DEC-2000 (first entry)

Cytomegalovirus US2-DL6 fusion protein.

US2; hCMV; major histocompatibility complex; MHC; class I; class II; antigen presentation; inhibition; CD8-positive; CD4-positive; T cell; transplant; gene therapy; immunosuppressive; fusion; DL6.

Chimeric - Human cytomegalovirus.
Chimeric - Apis sp.
Chimeric - Homo sapiens.
Chimeric - Synthetic.

Key location/Qualifiers
Peptide 1..21
Protein /label= Bee_mellitin_signal_peptide
Misc-difference 158 /label= Soluble_US2_protein
Peptide /note= "Single glycine spacer"
159..177 /label= DL6_epitope
WO200046361-A1.
10-AUG-2000.

Signal peptide of
Mellitin signal pe
Heterologous signa
Honeybee mellitin
Honeybee mellitin
Honeybee mellitin
Honeybee mellitin
Mellitin protein s
Honey bee mellitin
6His affinity tag
Human GIL-19/AE289
Baculovirus expres
Recombinant vaccin
HSV-1 US6-associat
Polypeptide with a
HSV gD1 peptide fr
HSV glycoprotein D
Herpes Simplex vir
Sequence encoded b
HSV-1 US6-associat
Honeybee mellitin
Mellitin signal se
Honey bee mellitin
Honey bee mellitin
HSV antigen gD2.
Sequence of Herpes
Herpes Simplex vir
HSV glycoprotein D
HSV-2 strain SB5 C
Sequence encoded b
HSV-2 strain SB5 C
HSV-2 strain SB5 C
PMelBac secretion
Targeting molecule
Secretion signal a

```
XX 02-FEB-2000; 2000WO-US02740.
PF
XX
XX 02-FEB-1999; 99US-0118287.
PR
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
PA
XX Johnson DC, Tomazin R, Boname J, Hegde NR;
PI WPI; 2000-506069/45.
DR N-PSDB; AAA53814.
XX
XX Inhibiting recognition of cellular tissue by CD8+ and CD4+ T cells, to
PT treat or prevent autoimmune diseases, and to improve gene therapy,
PT comprises introducing human cytomegalovirus US2 protein into cells
XX
XX Example 12; Page 49; 53pp; English.
PS
XX Human cytomegalovirus (hCMV) US2 protein, which has previously been shown
CC to block the major histocompatibility complex (MHC) class I antigen
CC presentation pathway, blocks the MHC class II pathway. The US2 protein
CC does not have to be mutated to cause inhibition of the MHC class II
CC pathway. The binding domain recognizes MHC I heavy chains, MHC II alpha
CC chains, and optionally DM-alpha chains. US2 has a double inhibitory
CC effect on the MHC class II pathway, inhibiting recognition of cellular
CC tissue by CD8-positive and CD4-positive T cells. US2 or its soluble
CC variants, can be used to reduce inappropriate immune responses. The US2
CC protein can be used to improve the persistence of a virus. Vectors
CC encoding soluble US2 protein (residues 28-143) can be used to treat
CC autoimmune disease, especially where it is mediated by MHC II molecules.
CC The vector can also be used to improve gene therapy, and preferably also
CC contains a sequence encoding a therapeutic product. The protein is
CC exogenously supplied or expressed from a recombinant cell, and
CC may also be used to inhibit CD4-positive mediated immune responses,
CC autoimmune responses, transplant immune responses and gene therapy
CC immune responses.
XX
XX Sequence 177 AA;
SQ
XX
XX Query Match 100.0%; Score 947; DB 21; Length 177;
Best Local Similarity 100.0%; Pred. No. 7.9e-99;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKRLVNVALVFMVYVYISYIYARLPDGTAKGEDALRPWKSTAKHPFQIEDNRCYINDGK 60
DB 1 MKRLVNVALVFMVYVYISYIYARLPDGTAKGEDALRPWKSTAKHPFQIEDNRCYINDGK 60
QY 61 LPARGSIVGNSRFVFDPKADYGVGENLYVHADDFEVPGESLKNVNRNLDMPIFETL 120
DB 61 LPARGSIVGNSRFVFDPKADYGVGENLYVHADDFEVPGESLKNVNRNLDMPIFETL 120
QY 121 ALRLVLQGDVIMLRVCYBELRDYTSAYMMNMQYGMVGQPELAPEDPEDSALLEDPV 177
DB 121 ALRLVLQGDVIMLRVCYBELRDYTSAYMMNMQYGMVGQPELAPEDPEDSALLEDPV 177
RESULT 2
AAAY97248
ID AAY97248 standard; Protein; 199 AA.
XX
XX AAY97248;
AC
XX 04-DEC-2000 (first entry)
DT
XX Cytomegalovirus US2 protein.
DE
XX
XX US2; hCMV; major histocompatibility complex; MHC; class I; class II;
KW antigen presentation; inhibition; CD8-positive; CD4-positive; T cell;
KW transplant; gene therapy; immunosuppressive.
XX
XX Human cytomegalovirus.
OS
XX WO200046361-A1.
XX
```

```
XX 10-AUG-2000.
PD
XX
XX 02-FEB-2000; 2000WO-US02740.
PF
XX
XX 02-FEB-1999; 99US-0118287.
PR
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
PA
XX Johnson DC, Tomazin R, Boname J, Hegde NR;
PI WPI; 2000-506069/45.
DR N-PSDB; AAA53813.
XX
XX Inhibiting recognition of cellular tissue by CD8+ and CD4+ T cells, to
PT treat or prevent autoimmune diseases, and to improve gene therapy,
PT comprises introducing human cytomegalovirus US2 protein into cells
XX
XX Claim 6; Page 47-48; 53pp; English.
PS
XX This is the human cytomegalovirus (hCMV) US2 protein, which has
CC previously been shown to block the major histocompatibility complex (MHC)
CC class I antigen presentation pathway, blocks the MHC class II pathway.
CC The US2 protein does not have to be mutated to cause inhibition of the
CC MHC class II pathway. The binding domain recognizes MHC I heavy chains,
CC MHC II alpha chains, and optionally DM-alpha chains. US2 has a double
CC inhibitory effect on the MHC class II pathway, inhibiting recognition of
CC cellular tissue by CD8-positive and CD4-positive T cells. US2 or its
CC soluble variants, can be used to reduce inappropriate immune responses.
CC The US2 protein can be used to improve the persistence of a virus.
CC Vectors encoding soluble US2 protein (residues 28-143) can be used to
CC treat autoimmune disease, especially where it is mediated by MHC II
CC molecules. The vector can also be used to improve gene therapy, and
CC preferably also contains a sequence encoding a therapeutic product. The
CC protein is exogenously supplied or expressed from a recombinant cell, and
CC may also be used to inhibit CD4-positive mediated immune responses,
CC autoimmune responses, transplant immune responses and gene therapy
CC immune responses.
XX
XX Sequence 199 AA;
SQ
XX
XX Query Match 78.1%; Score 740; DB 21; Length 199;
Best Local Similarity 100.0%; Pred. No. 2.5e-75;
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22 RLPDGTAKGEDALRPWKSTAKHPFQIEDNRCYINDGKLFPARGSIVGNSRFVFDPKAD 81
DB 20 RLPDGTAKGEDALRPWKSTAKHPFQIEDNRCYINDGKLFPARGSIVGNSRFVFDPKAD 79
QY 82 YGVGENLYVHADDFEVPGESLKNVNRNLDMPIFETLALRLVLQGDVIMLRVCYBELRV 141
DB 80 YGVGENLYVHADDFEVPGESLKNVNRNLDMPIFETLALRLVLQGDVIMLRVCYBELRV 139
QY 142 DYTSSAYMMNMQYGMV 157
DB 140 DYTSSAYMMNMQYGMV 155
RESULT 3
ABG31298
ID ABG31298 standard; Protein; 451 AA.
XX
XX ABG31298;
AC
XX 21-OCT-2002 (first entry)
DT
XX
XX Hydrophobic protein melittin tag tag-Human beta2-adrenergic receptor EE.
DE
XX
XX Signal sequence; ligand identification; hydrophobic protein;
KW transmembrane protein; monotopic membrane protein;
KW polytopic membrane protein; pump; channel; receptor kinase;
KW G protein-coupled receptor; transporter protein.
XX
```

```
OS Synthetic.
XX
XX WO200257792-A2.
XX
XX 25-JUL-2002.
XX
XX 19-DEC-2001; 2001WO-US50088.
XX
XX 29-DEC-2000; 2000US-258970P.
XX
XX (NEOG-) NEOGENESIS PHARM INC.
XX
XX Felsch JS, Annis DA, Kalghatgi K, Nash HM,
XX
XX WPI; 2002-599728/64.
XX
XX
XX Identifying ligand for hydrophobic protein based on affinity selection
XX PT which can operate in the presence of amphiphile without regard to the
XX PT specific biological function of hydrophobic target protein -
XX
XX Disclosure; Fig 2; 97pp; English.
XX
XX This invention relates to a novel method for identifying a ligand for a
XX CC hydrophobic protein. The method comprises selecting a ligand molecule by
XX CC affinity selection by exposing a hydrophobic target protein bound by an
XX CC amphiphile to a multiplicity of molecules to promote formation of at
XX CC least a complex between the hydrophobic target protein and the ligand
XX CC molecule, separating the complex from the unbound molecules, and
XX CC identifying the ligand molecule. The method of the invention is
XX CC useful for identifying a ligand for hydrophobic protein such as a
XX CC membrane, integral membrane, transmembrane, monotopic or polytopic
XX CC membrane, pump, channel, receptor kinase, G protein-coupled receptor,
XX CC or transporter protein, or membrane-associated enzyme, or Myc tag-EB
XX CC tag-human m2 mACHR, flag tag-human beta2 adrenergic receptor-EB tag,
XX CC human neurokinin 3 receptor-HSV tag-Myc tag, flag tag-human m1 mACHR-EB
XX CC tag, and rat m3 mACHR-HSV tag-Octahis tag. The ligand identified by the
XX CC method of the invention is useful for the development of novel
XX CC medicines and medicinal diagnostics. The present sequence represents
XX CC the hydrophobic protein melittin tag tag-Human beta2-adrenergic receptor
XX CC EB used in the method of the invention.
XX
XX SQ Sequence 451 AA;
XX
XX Query Match 12.4%; Score 117.5; DB 23; Length 451;
XX Best Local Similarity 27.7%; Pred. No. 0.00022;
XX Matches 41; Conservative 13; Mismatches 45; Indels 49; Gaps 3;
XX
XX QY 1 MKFLVNVALLVFMVVYISYIYARLPDGIITKAGEDALRPWKSTAKHPWFQIEDNRCYIDNGK 60
XX |||||
XX 1 MKFLVNVALLVFMVVYISYIYADYKDDDKMGQP----- 33
XX
XX QY 61 LPARGSIVGNMSRFVDPKADYGVGENLYHADVVEFVPGESLKMNVRLNDVMPFETL 120
XX |||||
XX 34 -----GNGSAFLAPNRS-----HAPDHVYTOQRDEWVVG---MGIVMSL 71
XX
XX QY 121 ALRLVLOGDVIMLRCPPELRVDYTSAY 148
XX : : : : :
XX Db 72 IVLAIVFGNVLVITAIKFERLQTVINY 99
XX
XX RESULT 4
XX ABG31300
XX ID ABG31300 standard; Protein; 498 AA.
XX
XX AC ABG31300;
XX
XX DT 21-OCT-2002 (first entry)
XX
XX DE Hydrophobic protein melittin-flag @tag-human m1 mACHR-EB.
XX
XX Signal sequence; ligand identification; hydrophobic protein;
XX KW transmembrane protein; monotopic membrane protein;
XX KW polytopic membrane protein; pump; channel; receptor kinase;
```

```
KW G protein-coupled receptor; transporter protein.
XX
XX OS Synthetic.
XX
XX PN WO200257792-A2.
XX
XX PD 25-JUL-2002.
XX
XX PF 19-DEC-2001; 2001WO-US50088.
XX
XX PR 29-DEC-2000; 2000US-258970P.
XX
XX (NEOG-) NEOGENESIS PHARM INC.
XX
XX PA Felsch JS, Annis DA, Kalghatgi K, Nash HM,
XX
XX PI WPI; 2002-599728/64.
XX
XX DR
XX
XX DR Identifying ligand for hydrophobic protein based on affinity selection
XX PT which can operate in the presence of amphiphile without regard to the
XX PT specific biological function of hydrophobic target protein -
XX
XX XX Disclosure; Fig 4; 97pp; English.
XX
XX This invention relates to a novel method for identifying a ligand for a
XX CC hydrophobic protein. The method comprises selecting a ligand molecule by
XX CC affinity selection by exposing a hydrophobic target protein bound by an
XX CC amphiphile to a multiplicity of molecules to promote formation of at
XX CC least a complex between the hydrophobic target protein and the ligand
XX CC molecule, separating the complex from the unbound molecules, and
XX CC identifying the ligand molecule. The method of the invention is
XX CC useful for identifying a ligand for hydrophobic protein such as a
XX CC membrane, integral membrane, transmembrane, monotopic or polytopic
XX CC membrane, pump, channel, receptor kinase, G protein-coupled receptor,
XX CC or transporter protein, or membrane-associated enzyme, or Myc tag-EB
XX CC tag-human m2 mACHR, flag tag-human beta2 adrenergic receptor-EB tag,
XX CC human neurokinin 3 receptor-HSV tag-Myc tag, flag tag-human m1 mACHR-EB
XX CC tag, and rat m3 mACHR-HSV tag-Octahis tag. The ligand identified by the
XX CC method of the invention is useful for the development of novel
XX CC medicines and medicinal diagnostics. The present sequence represents
XX CC the hydrophobic protein melittin-flag @tag-human m1 mACHR-EB used in the
XX CC method of the invention.
XX
XX SQ Sequence 498 AA;
XX
XX Query Match 11.6%; Score 110; DB 23; Length 498;
XX Best Local Similarity 51.8%; Pred. No. 0.0018;
XX Matches 29; Conservative 0; Mismatches 13; Indels 14; Gaps 2;
XX
XX QY 1 MKFLVNVALLVFMVVYISYIYARLPDGIITKAGEDA-----LRPWKSTAKHPW 46
XX |||||
XX Db 1 MKFLVNVALLVFMVVYISYIYADYKDDDKNTSAPPAVSPNITVLAP----GKGFW 52
XX
XX RESULT 5
XX AAM47008
XX ID AAM47008 standard; Protein; 1189 AA.
XX
XX AC AAM47008;
XX
XX DT 13-AUG-1998 (first entry)
XX
XX DE Glutathione-S-transferase and hTTRT fusion protein 8.
XX
XX Human; telomerase reverse transcriptase; hTTRT; TRT; diagnosis;
XX KW prognosis; cell proliferation; cancer; ageing; ribonucleoprotein.
XX
XX OS Synthetic.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 22..23 /note= "enterokinase cleavage site"
XX
XX FT
```

XX GB2317891-A.
PN
XX
PD 08-APR-1998.
XX
XX
PF 01-OCT-1997; 97GB-0020890.
XX
PR 14-AUG-1997; 97US-0915503.
PR 01-OCT-1996; 96US-0724643.
PR 18-APR-1997; 97US-0844419.
PR 25-APR-1997; 97US-0846017.
PR 06-MAY-1997; 97US-0851843.
PR 09-MAY-1997; 97US-0854050.
PR 14-AUG-1997; 97US-0911312.
PR 14-AUG-1997; 97US-0912951.
XX
XX
PA (GERO-) GERON CORP.
PA (UYTE-) UNIV TECHNOLOGY CORP.
XX
XX
PI Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J;
PI Morin GB, Nakamura T, Harley CB;
XX
XX
DR WPI, 1998-171633/16.
XX
XX
PT Pure and recombinant human Telomerase Reverse Transcriptase and its
PT variants - are useful in the diagnosis, prognosis and treatment of
PT cell proliferation conditions especially cancer and ageing
XX
PS Example 6; Page 234-235; 387pp; English.

The present sequence represents a fusion protein from an example of the present invention which describes human telomerase reverse transcriptase (hTERT). The present invention also describes the following methods: (A) determining whether a test compound is a modulator of hTERT, by detecting the change in hTERT recombinant protein or polynucleotide, on administration of the compound; (B) preparation of recombinant telomerase by contacting a protein preparation of hTERT with a telomerase RNA component; (C) detection of the hTERT RNA or protein in a sample by binding a relevant probe to the sample and detecting the complex formed or in the case of RNA detection, amplifying the product and correlating the presence of complex or amplification product with presence of hTERT in the sample; and (D) increasing the proliferation of a vertebrate cell by increasing hTERT expression; and (E) the use of an agent that causes an increase in cell vertebrate cell proliferation to create a medicament that inhibits ageing. A protein preparation of hTERT and the polynucleotide encoding hTERT can be used in the manufacture of medicaments for inhibiting the effect of ageing or cancer. Inhibitors of telomerase activity can be used to treat conditions that are associated with high telomerase activity. A protein preparation of hTERT can also be used in the new methods.

Query Match	11.6%;	Score 109.5;	DB 19;	length 1189;
Best Local Similarity	32.1%;	Pred. No. 0.0067;		
Matches	50;	Conservative 10;	Mismatches 65;	Indels 31; Gaps 6;
QY	1	MKFLVNVALVFVVVYIISYIYARLPDGITKAG--EDA-----	ILRPWKSTAKHPWFOI	49
DB	1	MKFLVNVALVFVVVYIISYIYAD-PSRSAGTMEFPAASTGRCVLLTWTWELLAPATPAMP		59
QY	50	EDNRCYIDNGKLFARGSIYGNMSREVPDPKADYGGVGENTLYHADVVEF-----V		99
DB	60	RAPRCRAVRSLSLRSHYREVLPLATV---RRLGPQGWRLLVORGDPAAFALVAQCLVCV		115
QY	100	PGESLKMVVRNLDVMPIFETLALRLVYQGDIWLMRC		135
DB	116	P-----WDARPPPAAPSFROVSCLELVARLYQRLC		146

RESULT 6
AAB10640

ID	AAB10640 standard; Protein; 354 AA.
XX	
XX	AAB10640;
AC	
XX	
DT	19-JAN-2001 (first entry)
XX	
DE	Human VEGF-X protein for expression in baculovirus/insect cell systems.
XX	
KW	VEGF-X; vascular endothelial growth factor; human; vulnery; cyostatic;
KW	antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
KW	angiogenesis regulator; vascularization regulator; cancer; psoriasis;
KW	rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
KW	tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KW	venous sore; diabetic ulcer; burns; skin graft growth.
XX	
OS	Homo sapiens.

PN	WO200037641-A2.
XX	
PD	29-JUN-2000.
XX	
PF	21-DEC-1999; 99WO-US30503.
XX	
PR	22-DEC-1998; 98GB-0028377.
PR	18-MAR-1999; 99US-0124967.
PR	08-NOV-1999; 99US-0164131.
XX	
PA	(JANC) JANSSEN PHARM NV.
XX	
PI	Gordon RD, Sprengel JJ, Yon JR, Dijkmans JH, Gosiewska A;
PI	Dhanaraj SN, Xu J;
XX	
DR	WPI; 2000-442669/38.
DR	N-PSDB; AAA71984.
XX	
PT	
PT	New vascular endothelial growth factor protein, useful for treating or
PT	preventing diseases associated with inappropriate angiogenesis activity
PT	such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX	
PS	Disclosure; Fig 20; 127pp; English.

This invention describes a novel vascular endothelial growth factor-X (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has vulnery, cytostatic, antirheumatic, antiarthritic, antipsoriatic and antidiabetic activity and acts as an angiogenesis and vascularization regulator. An antisense molecule of the invention is useful for treating or preventing cancer, rheumatoid arthritis, psoriasis and diabetic retinopathy by inhibiting angiogenic activity or inappropriate vascularization including formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair in a subject. The products of the invention are useful for preparing medicaments for treating wounds such as dermal ulcers, pressure sores, venous sores, diabetic ulcers and burns and to promote skin graft growth, tissue repair, proliferation of new blood vessels, tissue regeneration and organ repair by promoting angiogenic activity or vascularization. This sequence represents a human VEGF-X protein which can be expressed in Baculovirus/insect cell systems and which is described in the method of the invention.

```

Query Match          11.2%; Score 106.5; DB 21; Length 354;
Best Local Similarity 24.4%; Pred. No. 0.0028;
Matches 44; Conservative 14; Mismatches 41; Indels 81; Gaps 6;

QY 1 MKELVNVALLVFWVVYISYIYARLBDGITAGEDALRPWKSTPAKHPWQIEDNRCYIDNGK 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MKELVNVALLVFWVVYISYIYA-----DPESHHHHHHESNL----- 35

QY 61 LFARGSIVGNMSRFVFPDEKADYGVGE-----NLVYHADDFVFPGESLKNVVR 109
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 36 -----SSKFOFSNNKEONGVQDPQHERILITVSTNGSIHSP-----R 71

```

```
Qy 110 NLDVMEIFETLALRLVIGDVIWLRCPPELRVDYTSSAYMMNQYGMVGPPELADPEP 169
Db 72 FPHYPRNTVLVWRLVAEENWVLIQT-----FDERFGL-----EDPED 110
RESULT 7
AAB10641
ID AAB10641 standard; Protein; 354 AA.
AC AAB10641;
XX
XX 19-JAN-2001 (first entry)
DE Human VEGF-X protein for expression in E. coli systems.
XX
XX VEGF-X; vascular endothelial growth factor; human; vulnery; cyostatic;
KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KW venous sore; diabetic ulcer; burns; skin graft growth.
XX
XX Homo sapiens.
XX
XX WO200037641-A2.
XX
XX 29-JUN-2000.
XX
XX 21-DEC-1999; 99WO-US30503.
XX
XX 22-DEC-1998; 98GB-0028377.
XX
XX 18-MAR-1999; 99US-0124967.
XX
XX 08-NOV-1999; 99US-0164131.
XX
XX (JANC ) JANSSEN PHARM NV.
PA
XX
XX Gordon RD, Sprengel J, Yon JR, Dijkmans JH, Gosiowska A;
PI Dhanaraj SN, Xu J;
PI
XX
XX WPI; 2000-442669/38.
XX
XX N-PSDB; AAA71985.
XX
XX New vascular endothelial growth factor protein, useful for treating or
XX preventing diseases associated with inappropriate angiogenesis activity
XX PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX
XX Disclosure; Fig 21; 127pp; English.
XX
XX This invention describes a novel vascular endothelial growth factor-X
XX (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
XX vulnery, cyostatic, antirheumatic, antiarthritic, antipsoriatic and
XX antidiabetic activity and acts as an angiogenesis and vascularization
XX regulator. An antisense molecule of the invention is useful for treating
XX or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
XX retinopathy by inhibiting angiogenic activity or inappropriate
XX vascularization including formation and proliferation of new blood
XX vessels, growth and development of tissues, tissue regeneration and organ
XX and tissue repair in a subject. The products of the invention are useful
XX for preparing medicaments for treating wounds such as dermal ulcers,
XX pressure sores, venous sores, diabetic ulcers and burns and to promote
XX skin graft growth, tissue repair, proliferation of new blood vessels,
XX tissue regeneration and organ repair by promoting angiogenic activity or
XX vascularization. This sequence represents a human VEGF-X protein which
XX can be expressed in E. coli systems and which is described in the method
XX of the invention.
XX
XX Sequence 354 AA;
SQ
Query Match 11.2%; Score 106.5; DB 21; Length 354;
Best Local Similarity 24.4%; Pred. No. 0.0028;
Matches 44; Conservative 14; Mismatches 41; Indels 81; Gaps 6;
```

```
Db 1 MKFLVNVALVFMVYISYIYARLPDGIITKAGEDALPWXSTAKHPWFQIEDNRCYIDNGK 60
Qy 110 NLDVMEIFETLALRLVIGDVIWLRCPPELRVDYTSSAYMMNQYGMVGPPELADPEP 169
Db 72 FPHYPRNTVLVWRLVAEENWVLIQT-----FDERFGL-----EDPED 110
RESULT 8
AAE05128
ID AAE05128 standard; Protein; 386 AA.
XX
XX AAE05128;
AC
XX 18-SEP-2001 (first entry)
DT
XX Infectious salmon anaemia virus (ISAV) 92-M fusion protein.
XX
XX Infectious salmon anaemia virus; ISAV; 92-M clone; vaccine; prophylaxis;
KW infectious salmon anaemia; therapy; fusion protein; antianaemic.
XX
XX Chimeric - Infectious salmon anaemia virus.
XX
XX Chimeric - Unidentified.
XX
XX Key Location/Qualifiers
XX FT 1..21
XX FT /label= Melittin_signal_peptide
XX FT 22..26
XX FT /label= linker_peptide
XX FT 27..386
XX FT /note= "Mature ISAV 92-M protein"
XX
XX WO200149712-A2.
XX
XX 12-JUL-2001.
XX
XX 03-JAN-2001; 2001WO-EP00046.
XX
XX 07-JAN-2000; 2000EP-020054.
XX
XX 29-FEB-2000; 2000EP-0200700.
XX
XX (ALKU ) AKZO NOBEL NV.
XX
XX Biering E, Krossoy B;
PI
XX
XX WPI; 2001-441845/47.
XX
XX N-PSDB; AAD09875.
XX
XX Novel vaccine for treatment and/or prevention of infectious salmon
XX anemia in fish -
XX
XX Claim 7; Page 39-41; 43pp; English.
XX
XX The present invention relates to vaccine for prevention and/or
XX prophylaxis of infectious salmon anaemia in fish. The invention
XX provides for nucleic acid sequence encoding viral proteins of
XX infectious salmon anaemia virus (ISAV) as well as the isolated protein.
XX Nucleic acids encoding viral protein is useful for the manufacture of a
XX DNA vaccine for diagnosis, treatment and/or prophylaxis of infectious
XX salmon anaemia in fish, and viral protein can be used for the
XX manufacture of antibodies that are specific for ISAV. The present
XX sequence is infectious salmon anaemia virus 92-M fusion protein.
XX
XX Sequence 386 AA;
SQ
Query Match 11.0%; Score 104; DB 22; Length 386;
Best Local Similarity 59.0%; Pred. No. 0.006;
Matches 23; Conservative 5; Mismatches 9; Indels 2; Gaps 1;
```

OY 1 MKFLVNVALVFMVVYISYIYARLPD--GITKAGEDALRP 37
 DB 1 MKFLVNVALVFMVVYISYIAEFKGLMGDSRSDQSRVNP 39

RESULT 9

ID ABG31502 standard; Protein; 496 AA.

ABG31502;

21-OCT-2002 (first entry)

Muscarinic acetylcholine receptor protein in Baculovirus expression.

Signal sequence; ligand identification; hydrophobic protein;
 transmembrane protein; monotopic membrane protein;
 polytopic membrane protein; pump; channel; receptor kinase;
 G protein-coupled receptor; transporter protein.

Synthetic.

WO200257792-A2.

25-JUL-2002.

19-DEC-2001; 2001WO-US50088.

29-DEC-2000; 2000US-258970P.

(NEOG-) NEOGENESIS PHARM INC.

Felisch JS, Annis DA, Kalghatgi K, Nash HM;

WPI; 2002-599728/64.

Identifying ligand for hydrophobic protein based on affinity selection
 PT which can operate in the presence of amphiphile without regard to the
 PT specific biological function of hydrophobic target protein -

Example 2; Page 45; 97pp; English.

This invention relates to a novel method for identifying a ligand for a
 CC hydrophobic protein. The method comprises selecting a ligand molecule by a
 CC affinity selection by exposing a hydrophobic target protein bound by an
 CC amphiphile to a multiplicity of molecules to promote formation of at
 CC least a complex between the hydrophobic target protein and the ligand
 CC molecule, separating the complex from the unbound molecules, and
 CC identifying the ligand molecule. The method of the invention is
 CC useful for identifying a ligand for hydrophobic protein such as a
 CC membrane, integral membrane, transmembrane, monotopic or polytopic
 CC membrane, pump, channel, receptor kinase, G protein-coupled receptor,
 CC or transporter protein, or membrane-associated enzyme, or Myc tag-EB
 CC tag-human m2 mACHR, flag tag-human beta2 adrenergic receptor-EB tag,
 CC human neurokinin 3 receptor-HSV tag-Myc tag, flag tag-human m1 mACHR-EB
 CC tag, and rat m3 mACHR-HSV tag-Octahis tag. The ligand identified by the
 CC method of the invention is useful for the development of novel
 CC medicines and medicinal diagnostics. The present sequence represents
 CC the protein sequence of muscarinic acetylcholine receptor in a
 CC baculovirus expression vector used in the method of the invention.

Sequence 496 AA;

Query Match 11.0%; Score 104; DB 23; Length 496;

Best Local Similarity 88.0%; Pred. No. 0.0084;

Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 MKFLVNVALVFMVVYISYIYARLPD 25
 DB 1 MKFLVNVALVFMVVYISYIYADYKD 25

RESULT 10

AAG63690
 ID AAG63690 standard; peptide; 21 AA.
 AC AAG63690;

29-OCT-2001 (first entry)

Signal peptide of bee venom melittin.

Active polypeptide; ectoparasitic leech; Rhynchobdellida; C3bB complex;
 complement activation; complement factor D; haemodialysis; catheter;
 cardio-pulmonary bypass; extra-arterial stent; transplant rejection;
 autoimmune disease; lupus arthritis; rheumatoid arthritis; sepsis;
 glomerulonephritis; nephritis; nephropathy; reperfusion; anaphylaxis;
 asthma; skin reaction; infection; sickle cell anemia; haemolytic anemia.

Apis sp.

WO200147963-A2.

05-JUL-2001.

21-DEC-2000; 2000WO-GB04971.

24-DEC-1999; 99GB-0030659.

(BIOD-) BIO-DISCOVERY LTD.

Finney S, Seale L, Wallis RB;

WPI; 2001-522011/57.

Novel polypeptides from the leech Placobdella papillifera which inhibit
 PT alternate pathway of complement activation, useful for treating
 PT rheumatoid arthritis, sepsis, asthma involving alternate complement
 PT pathway -

Disclosure; Page 14; 80pp; English.

The present sequence represents a signal peptide, which may be used
 CC to express leech active polypeptides. The active polypeptide has
 CC a molecular weight of 7000-17,000 Da (as measured by mass
 CC spectrometry), and is derived from ectoparasitic leeches, of order
 CC Rhynchobdellida, of genus Placobdella and especially of species
 CC P. papillifera. The polypeptide inhibits the alternate route of
 CC complement activation but has substantially no effect on complement
 CC activation by the classical route. The polypeptide is an inhibitor of
 CC complement factor D and/or C3bB complex. The active polypeptide is
 CC useful for manufacturing a medicament and inhibits one or more steps
 CC in the alternate pathway of complement activation. It is useful for
 CC treating or preventing conditions, such as haemodialysis and
 CC cardio-pulmonary bypass, the presence of in-dwelling catheters and
 CC extra-arterial stents, rejection of transplanted organs or tissues,
 CC autoimmune diseases including lupus arthritis, rheumatoid arthritis,
 CC glomerulonephritis, nephritis, nephropathy, sepsis, injury caused to
 CC tissues by reperfusion after an ischaemic period and other conditions
 CC associated with activation of complement, including anaphylaxis,
 CC asthma, skin reactions, infections, sickle cell anemia and haemolytic
 CC anemia involving activation of alternate complement pathway in a
 CC patient.

Sequence 21 AA;

Query Match 10.8%; Score 102; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 0.00019;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKFLVNVALVFMVVYISYIYA 21
 DB 1 MKFLVNVALVFMVVYISYIYA 21

RESULT 11

AAB72437
ID AAB72437 standard; Peptide; 21 AA.
XX
AC AAB72437;
XX
DT 08-MAY-2001 (first entry)
XX
DE Melittin signal peptide.
XX
KW UGCT; UDP-glucose:glycoprotein glucosyltransferase; enzyme; honeybee;
KW melittin signal peptide.
XX
OS Apis mellifica.
XX
PN WO200112845-A1.
XX
PD 22-FEB-2001.
XX
PF 27-JUL-2000; 2000WO-CA00883.
XX
PR 18-AUG-1999; 99US-0376330.
XX
PA (CANA) NAT RES COUNCIL CANADA.
XX
PI Tessier DC, Dignard D, Bergeron JJM, Thomas DY;
XX
DR WPI; 2001-218358/22.
XX
DR N-PSDB; AAF60733.
XX
PT Determining the effect of a test sample on UDP-glucose:glycoprotein
PT glucosyltransferase (UGCT), useful for measuring UGCT activity,
PT comprises exposing an acceptor substrate for UGCT to a labeled donor in
PT the presence of UGCT -
XX
XX Example 1; Fig 8; 95pp; English.
XX
PS The present invention relates to a method for determining the effect of a
CC test sample on UDP-glucose:glycoprotein glucosyltransferase (UGCT)
CC activity. The method comprises exposing an acceptor substrate for UGCT to
CC a labelled donor in the presence of the test sample and UGCT. The method
CC is useful for determining UGCT activity. In particular, the method is
CC useful in glucosyltransferase assay and kinetics measurement for
CC determining UGCT activity. UGCT is a soluble enzyme of the endoplasmic
CC reticulum which catalyses the addition of a glucose residue onto
CC asparagine-linked oligosaccharides, which are present on incorrectly
CC folded glycoproteins. The present sequence is the honeybee melittin
CC signal peptide. This sequence was used in the construction of an
CC expression vector for rat UGCT (see AAF60732 and AAB72436).
XX
XX
SQ Sequence 21 AA;
QY
Query Match 10.8%; Score 102; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 MKFLVNVALVFMVVYISYIYA 21
1 MKFLVNVALVFMVVYISYIYA 21

RESULT 12
ABG31292
ID ABG31292 standard; Peptide; 21 AA.
XX
AC ABG31292;
XX
DT 21-OCT-2002 (first entry)
XX
DE Heterologous signal sequence #1.
XX
KW Signal sequence; ligand identification; hydrophobic protein;
KW transmembrane protein; monotopic membrane protein;
KW polytopic membrane protein; pump; channel; receptor kinase;

KW G protein-coupled receptor; transporter protein.
XX
XX Synthetic.
OS
XX
PN WO200257792-A2.
XX
PD 25-JUL-2002.
XX
PF 19-DEC-2001; 2001WO-US50088.
XX
PR 29-DEC-2000; 2000US-258970P.
XX
PA (NEOG-) NEOGENESIS PHARM INC.
XX
PI Felsch JS, Annis DA, Kalghatgi K, Nash HM,
XX
DR WPI; 2002-599728/64.
XX
PT Identifying ligand for hydrophobic protein based on affinity selection
PT which can operate in the presence of amphiphile without regard to the
PT specific biological function of hydrophobic target protein -
XX
XX Claim 32; Page 80; 97pp; English.
XX
PS This invention relates to a novel method for identifying a ligand for a
CC hydrophobic protein. The method comprises selecting a ligand molecule by
CC affinity selection by exposing a hydrophobic target protein bound by an
CC amphiphile to a multiplicity of molecules to promote formation of at
CC least a complex between the hydrophobic target protein and the ligand
CC molecule, separating the complex from the unbound molecules, and
CC identifying the ligand molecule. The method of the invention is
CC useful for identifying a ligand for hydrophobic protein such as a
CC membrane, integral membrane, transmembrane, monotopic or polytopic
CC membrane, pump, channel, receptor kinase, G protein-coupled receptor,
CC or transporter protein, or membrane-associated enzyme, or Myc tag-EE
CC tag-human m2 MACHR, flag tag-human beta2 adrenergic receptor-BE tag,
CC human neurokinin 3 receptor-HSV tag-Myc tag, flag tag-human m1 MACHR-BE
CC tag, and rat m3 MACHR-HSV tag-Octahis tag. The ligand identified by the
CC method of the invention is useful for the development of novel
CC medicines and medical diagnostics. The present sequence represents
CC a heterologous signal sequence peptide used in the method of the
CC invention.
XX
XX
SQ Sequence 21 AA;
QY
Query Match 10.8%; Score 102; DB 23; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 MKFLVNVALVFMVVYISYIYA 21
1 MKFLVNVALVFMVVYISYIYA 21

RESULT 13
ABB10034
ID ABB10034 standard; Protein; 21 AA.
XX
AC ABB10034;
XX
DT 08-OCT-2002 (first entry)
XX
DE Honeybee melittin signal sequence.
XX
KW Honeybee; melittin; signal sequence; immunoglobulin; B-cell lymphoma;
KW B-cell mediated pathology; cytostatic; immunosuppressive; dermatological;
KW antiinflammatory; neuroprotective; antidiabetic; antithyroid;
KW autoimmune disease.
XX
XX Apis mellifica.
OS
XX
PN WO200213862-A2.
XX

PD 21-FEB-2002.
XX
XX 10-AUG-2001; 2001WO-US25204.
PF
XX 11-AUG-2000; 2000US-224722P.
PR 11-AUG-2000; 2000US-224723P.
PR 23-MAR-2001; 2001US-279079P.
XX
PA (FAVR-) FAVRILLE INC.
XX
PI Gold DP, Shopes RJ;
XX WPI; 2002-280742/32.
DR N-PSDB; ABL54969.
XX
PT Composition for altering B-cell mediated pathology, has a chimeric
PT protein having portion of variable region of heavy chain or light chain
PT linked to portion constant region associated with patient B cell clone
PT
PS Claim 19; Page 17; 100pp; English.
XX
XX The sequence represents the honeybee melittin signal sequence, used in
CC the invention to maximise secretion of heterologous proteins from insect
CC cells. The invention relates to a novel composition for altering a B-cell
CC mediated pathology in a patient. The composition contains a chimeric
CC protein comprising at least a portion of a variable region of heavy chain
CC or light chain (VH or VL) linked to at least a portion of an
CC immunoglobulin constant region, where VH or VL region is associated with
CC a B cell clone from the patient having the B cell mediated pathology. The
CC composition of the invention has cytostatic, immunosuppressive,
CC dermatological, antiinflammatory, neuroprotective, antidiabetic, and
CC antithyroid activity. The composition is a vaccine useful for altering a
CC B cell mediated pathology. This includes B cell lymphoma e.g.
CC non-Hodgkins lymphoma, refractory low grade or follicular B cell
CC lymphoma; autoimmune disease e.g. multiple sclerosis, systemic lupus
CC erythematosus, anti-Hu associated paraneoplastic neurological syndrome,
CC autoimmune hepatitis, Type I diabetes, autoimmune thyroiditis and
CC scleroderma. The pathology is treated by administering the composition to
CC the patient, preferably with a cytokine e.g. granulocyte-macrophage-
CC colony stimulating factor (GM-CSF) or chemokine e.g. monocyte chemoattract
CC protein 3 (MCR 3).
XX
SQ Sequence 21 AA;
XX
Query Match 10.8%; Score 102; DB 23; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKFLVNVALVFMVYISYIYA 21
DB 1 MKFLVNVALVFMVYISYIYA 21
RESULT 14
AAU76918
ID AAU76918 standard; Peptide; 21 AA.
XX
XX AAU76918;
AC
XX
DT 05-JUN-2002 (first entry)
XX
XX Honeybee melittin signal sequence.
DE
XX Honeybee; signal sequence; rheumatoid arthritis; melittin;
KW Vbeta; Valpha; T cell receptor; TCR; cytosstatic; neuroprotective;
KW antidiabetic; dermatological; antirheumatic; immunosuppressive;
KW antiinflammatory; antiarthritic; antithyroid; T cell lymphoma;
KW autoimmune disease; multiple sclerosis; systemic lupus erythematosus;
KW diabetes; inflammatory bowel disease; myasthenia gravis; thyroiditis.
XX
XX Apis mellifica.
OS
XX

PN WO200213861-A2.
XX
XX 21-FEB-2002.
PD
XX 10-AUG-2001; 2001WO-US25203.
PF
XX 11-AUG-2000; 2000US-224722P.
PR 11-AUG-2000; 2000US-224723P.
PR 01-FEB-2001; 2001US-266133P.
XX
XX (FAVR-) FAVRILLE INC.
PA
XX Gold DP, Shopes RJ;
PI
XX WPI; 2002-241838/29.
DR N-PSDB; ABK10526.
XX
XX Altering T cell mediated pathology in patient, involves administering
PT composition having chimeric protein which has portion of Vbeta/Valpha
PT region of T cell receptor and portion of immunoglobulin constant region
PT
PS Disclosure; Page 8; 110pp; English.
XX
XX This invention relates to a method for altering a T cell mediated
CC pathology in a patient, the method involves administering a composition
CC comprising a chimeric protein which comprises at least a portion of a
CC Vbeta or Valpha region of a T cell receptor (TCR), and at least a
CC portion of an immunoglobulin (Ig) constant region, where the Vbeta or
CC Valpha region is associated with a particular TCR from a T cell from the
CC patient having T cell mediated pathology. The method may have
CC cytostatic, neuroprotective, antidiabetic, dermatological,
CC antirheumatic, immunosuppressive, antiinflammatory, antiarthritic and
CC antithyroid activities. The composition of the invention is useful for
CC altering a T cell mediated pathology in a patient, where the T cell
CC mediated pathology is T cell lymphoma, or an autoimmune disease selected
CC from multiple sclerosis, systemic lupus erythematosus, diabetes,
CC inflammatory bowel disease, myasthenia gravis, rheumatoid arthritis, and
CC thyroiditis. The present sequence represents the honeybee melittin
CC secretory signal sequence, this sequence was used as a leader sequence
CC in to maximise the secretion of heterologous proteins from transfected
CC insect cells.
XX
SQ Sequence 21 AA;
XX
Query Match 10.8%; Score 102; DB 23; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKFLVNVALVFMVYISYIYA 21
DB 1 MKFLVNVALVFMVYISYIYA 21
RESULT 15
AAM52328
ID AAM52328 standard; Peptide; 21 AA.
XX
XX AAM52328;
AC
XX
DT 28-JAN-2002 (first entry)
XX
XX Honeybee melittin signal peptide #1.
DE
XX Honeybee; signal peptide; melittin.
KW
XX Apis mellifica.
OS
XX JP2001258565-A.
PN
XX 25-SEP-2001.
PD
XX 21-MAR-2000; 2000JP-0078897.
PF

XX 21-MAR-2000; 2000JP-0078897.
PR (HARA/) HARA T.
XX
PA
XX WPI; 2002-003332/01.
DR N-PSDB; AAI71796.
XX
PT A nucleic acid for promoting the expression of genes, comprises a
PT secretion signal from chicken lysozyme or Sf162 type human
PT immunodeficiency syndrome virus -
XX
PS Example 1; Page 14; 21pp; Japanese.
XX
CC The present invention relates to an expression cassette for executing or
CC promoting the expression of a structural gene. The expression cassette
CC comprises a promoter connected to said structural gene and a secretion
CC signal peptide which is arranged upstream of the structural gene. The
CC present sequence is a signal peptide, which was used in the present
CC invention.
CC Note: the present sequence is the SEQ ID 8 as given in the sequence
CC listing. This sequence differs from the SEQ ID 8 given on page 8 (see
CC AAM52506).
XX
SQ Sequence 21 AA;

Query Match 10.8%; Score 102; DB 23; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKFLVNVALLVFMVVYISYIYA 21
Db 1 MKFLVNVALLVFMVVYISYIYA 21

Search completed: January 22, 2004, 11:44:40
Job time : 65.6382 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2004, 11:37:25 ; Search time 26.5802 Seconds
(without alignments)
313.155 Million cell updates/sec

Title: US-09-890-806-5

Perfect score: 947
Sequence: 1 MKFLVNVALVFMVVYISYIY.....GQPELAPDEPDSALLDPV 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	740	78.1	199	1	US02_HCMVA	P09713 human cytom
2	125	13.2	186	1	US03_HCMVA	P09712 human cytom
3	102	10.8	70	1	MEL_APIME	P01501 apis mellif
4	102	10.8	70	1	MEL_VESMC	P59262 vespula mac
5	99	10.5	393	1	VGLD_HSVIH	P06476 herpes simp
6	99	10.5	394	1	VGLD_HSVI1	O69031 herpes simp
7	99	10.5	394	1	VGLD_HSV1A	P36318 herpes simp
8	99	10.5	394	1	VGLD_HSV1P	P57083 herpes simp
9	98.5	10.4	70	1	MEL_APICC	P59260 apis cerana
10	98.5	10.4	70	1	MEL_POLAE	P59261 polistes he
11	96.5	10.2	1409	1	COP1_DROME	P04146 drosophila
12	91	9.6	393	1	VGLD_HSV2	P03172 herpes simp
13	81	8.6	410	1	REP1_ZYGRO	P13778 zygosacchar
14	80	8.4	703	1	CDGT_BACS2	P31746 bacillus sp
15	78.5	8.3	368	1	YJHT_ECOLI	P39371 escherichia
16	77.5	8.2	248	1	GRAN_GVXN	O12705 xestia c-ni
17	77	8.1	272	1	GPH1_PSEAE	O98586 pseudomonas
18	76.5	8.1	248	1	GRAN_GVTN	P06503 trichoplusi
19	76.5	8.1	316	1	AKBA_HUMAN	O60218 homo sapien
20	75	7.9	315	1	ALDR_PIG	P80276 sus scrofa
21	75	7.9	620	1	BCHD_CHLTE	O935W0 chlorobium
22	74.5	7.9	248	1	GRAN_GVAO	O91B74 adoxophyes
23	74	7.8	261	1	ZNUB_ECOLI	P39832 escherichia
24	73	7.7	463	1	UHPT_SALTY	P27670 salmonella
25	73	7.7	1620	1	HEDI_HUMAN	O9UL8 homo sapien
26	72.5	7.7	446	1	LAMB_ECOL6	O8CV14 escherichia
27	72.5	7.7	466	1	HEMN_SYNY3	P74132 synechocyst
28	72.5	7.7	619	1	BCHD_CHLVI	O50313 chlorobium
29	72	7.6	192	1	VIF_HVIB5	P04598 human immun
30	72	7.6	704	1	CDGT_BACOH	P27036 bacillus oh
31	71.5	7.6	1036	1	ATHL_RAT	P54708 rattus norv
32	71.5	7.6	1266	1	NGCA_CHICK	O03696 gallus gall
33	71	7.5	308	1	AMID_STREP	P18794 streptococc

34	71	7.5	491	1	PBP_BACSU	P39844 bacillus su
35	70.5	7.4	361	1	GLNA_PANAR	O04831 panulirus a
36	70.5	7.4	607	1	LEPA_STREP	O97qk5 streptococc
37	70.5	7.4	811	1	MUTS_THBAQ	O56215 thermus aqu
38	70	7.4	446	1	LAMB_ECOLI	P02943 escherichia
39	70	7.4	859	1	OBP_HSVBC	P52377 bovine hearp
40	69.5	7.3	349	1	US30_HCMVA	P09706 human cytom
41	69.5	7.3	446	1	LAMB_ECO57	O8X5W7 escherichia
42	69.5	7.3	562	1	GUNI_ACICE	P54583 acidothermu
43	69.5	7.3	878	1	SVV_METUA	O58413 methanococc
44	69.5	7.3	1059	1	CERV_RAT	P13635 rattus norv
45	69	7.3	286	1	ATPS_SCHPO	P87127 schizosacch

ALIGNMENTS

RESULT 1	US02_HCMVA	STANDARD;	PRT;	199 AA.
ID	US02_HCMVA			
AC	P09713;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Hypothetical protein HQLF2.			
GN	US2.			
OS	Human cytomegalovirus (strain AD169).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Betaherpesvirinae; Cytomegalovirus.			
ON	NCBI_Taxid=10360;			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=87169717; PubMed=3031311;			
RA	Weston K., Barrell B.G.;			
RT	"Sequence of the short unique region, short repeats, and part of the			
RT	long repeats of human cytomegalovirus.";			
RL	J. Mol. Biol. 192:177-208(1986).			
RN	[2]			
RP	COMPLETE GENOME.			
RX	MEDLINE=90269039; PubMed=2161319;			
RA	Chee M.S., Bankier A.T., Beck S., Bohm R., Brown C.M., Cerny R.,			
RA	Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,			
RA	Predlie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;			
RT	"Analysis of the protein-coding content of the sequence of human			
RT	cytomegalovirus strain AD169.";			
RL	Curr. Top. Microbiol. Immunol. 154:125-169(1990).			
CC	-1- SIMILARITY: BELONGS TO THE US2 FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X17403; CAA35313.1; -.			
DR	EMBL; X04650; CAB37096.1; -.			
DR	PIR; E26078; QOBECS.			
DR	PDB; 1IM3; 27-JUN-01.			
KW	Hypothetical protein; 3D-structure.			
FT	CARBOHYD 68			
FT	CARBOHYD 68			
FT	CARBOHYD 172			
FT	CARBOHYD 188			
FT	SEQUENCE 199 AA; 23111 MW; 4ADD2DF3D692393F3 CRC64;			
QY	Query Match	78.1%;	Score 740;	DB 1;
QY	Best Local Similarity	100.0%;	Pred. No. 3e-65;	Length 199;
QY	Matches 136;	Conservative 0;	Mismatches 0;	Indels 0;
QY				Gaps 0;
Db	22 RLPDGTAKGADALRPMKSTAKHPQIEDNRCYIDNGKLFARGSIGVNMRSFVDPKAD 81			
Db	20 RLPDGTAKGADALRPMKSTAKHPQIEDNRCYIDNGKLFARGSIGVNMRSFVDPKAD 79			


```

RN [8]
RP REVIEW.
RX MEDLINE=90254148; PubMed=2187536;
RA Dempsey C.E.;
RT "The actions of melittin on membranes.";
RL Biochim. Biophys. Acta 1031:143-161(1990).
CC -1- FUNCTION: Main toxin of bee venom with strong hemolytic activity.
CC integrates into cell membranes and has multiple effects, probably,
CC as a result of its interaction with negatively charged
CC phospholipids. It inhibits well known transport pumps such as the
CC Na(+)-K(+)-ATPase and the H(+)-K(+)-ATPase. Increases the
CC permeability of cell membranes to ions, particularly Na+ and
CC indirectly Ca(2+), because of the Na(+)-Ca(2+)-exchange.
CC -1- SUBUNIT: Monomer and homotetramer.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- MISCELLANEOUS: N-formyl-melittin major has 80% of the activity of
CC melittin.
CC -1- SIMILARITY: BELONGS TO THE MELLITIN FAMILY.
CC -1- DATABASE: NAME=Protein Spotlight;
CC NOTE=Issue 12 of July 2001;
CC WWW="http://www.expasy.org/spotlight/articles/sptlt012.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02007; CA26038.1; -.
DR PIR; A01762; MERHBI.
DR PIR; A91133; MERHBI.
DR PDB; 2MLT; 15-JUL-92.
DR PDB; 1BHL; 06-JAN-99.
DR InterPro; IPR002116; Melittin.
DR Pfam; PF01372; Melittin; 1.
DR ProDom; PD014636; Melittin; 1.
DR Cytolsys; Hemolysis; Toxin; Allergen; signal; Amidation; Formylation;
DR 3D-structure.
KW SIGNAL 1 21
FT PROPEP 22 43 REMOVED BY A DIPEPTIDYLPEPTIDASE.
FT CHAIN 44 69 MELLITIN.
FT MOD_RES 44 44 FORMYLATION (ABOUT 10% OF THE MOLECULES).
FT MOD_RES 69 69 AMIDATION (G-70 PROVIDE AMIDE GROUP).
FT VARIANT 64 64 K -> S (IN MELLITIN 2; POSSIBLY AN
FT ARTIFACT).
FT VARIANT 67 70 ROQG -> KRQG (IN MELLITIN 2; POSSIBLY AN
FT ARTIFACT).
FT HELIX 45 53
FT TURN 54 54
FT HELIX 55 68
SQ SEQUENCE 70 AA; 7585 MW; 607F52C091C23BB6 CRC64;
Query Match 10.8%; Score 102; DB 1; Length 70;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MKFLVNVALVFVYISYIYA 21
DB 1 MKFLVNVALVFVYISYIYA 21
RESULT 4
MEL_VESMC STANDARD; PRT; 70 AA.
AC P59262;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Melittin precursor.
GN MELT.

```

```

OS Vespa maculifrons (Eastern yellow jacket) (Wasp).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Vespoidea;
OC Vespidae; Vespinae; Vespa.
OX NCBI_TaxID=7453;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Shi W.J., Zhang S.F., Cheng J.A.;
RT "Cloning and sequencing of cDNA coding for prepromelittin of Polistes
RT hebraeus, Vespa magnifica and Vespa maculifrons.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Has strong hemolytic activity. Integrates into cell
CC membranes and has multiple effects, probably, as a result of its
CC interaction with negatively charged phospholipids. It inhibits
CC well known transport pumps such as the Na(+)-K(+)-ATPase and the
CC H(+)-K(+)-ATPase. Increases the permeability of cell membranes to
CC ions, particularly Na+ and indirectly Ca(2+), because of the
CC Na(+)-Ca(2+)-exchange (By similarity).
CC -1- SUBUNIT: Monomer and homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- SIMILARITY: BELONGS TO THE MELLITIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF487911; AA012205.1; -.
DR Pfam; PF01372; Melittin; 1.
DR ProDom; PD014636; Melittin; 1.
DR Cytolsys; Hemolysis; Toxin; signal; Amidation.
FT SIGNAL 1 21 BY SIMILARITY.
FT PROPEP 22 43 REMOVED BY A DIPEPTIDYLPEPTIDASE (BY
FT SIMILARITY).
FT CHAIN 44 69 MELLITIN.
FT MOD_RES 69 69 AMIDATION (G-70 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
SQ SEQUENCE 70 AA; 7585 MW; 607F52C091C23BB6 CRC64;
Query Match 10.8%; Score 102; DB 1; Length 70;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MKFLVNVALVFVYISYIYA 21
DB 1 MKFLVNVALVFVYISYIYA 21
RESULT 5
VGLD_HSV1H STANDARD; PRT; 393 AA.
AC P06476;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein D precursor.
GN GD OR US6.
OS Herpes simplex virus (type 1 / strain HZT).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10305;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84131549; PubMed=6321120;
RA Laskey L.A., Dowdenko D.J.;
RT "DNA sequence analysis of the type-common glycoprotein-D genes of
RT herpes simplex virus types 1 and 2.";
RL DNA 3:23-29(1984).

```

```

CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN D FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, K02372; AAA45786.1; -.
DR InterPro, IPR002896; Herpes_glycop_D.
DR Pfam, PF01537; Herpes_glycop_D; 1.
KT Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 393 GLYCOPROTEIN D.
FT DOMAIN 26 338 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 339 363 POTENTIAL.
FT DOMAIN 364 393 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 364 380 ARG/LYS-RICH (HIGHLY BASIC, PROBABLY
FT SERVES TO ANCHOR THE GLYCOPROTEIN IN
FT THE MEMBRANE).
FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 393 AA; 43368 MW; 348E8D35BA4B6FE9 CRC64;

```

Query Match	10.5%;	Score 99;	DB 1;	Length 393;
Best Local Similarity	100.0%;	Pred. No. 0.032;		
Matches 19;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      159 QPELAPEDPEDSALLDPV 177
        |||||
Db      289 QPELAPEDPEDSALLDPV 307
```

```

RESULT 6
ID      _VGLD_HSV1.1          STANDARD;          PRT;          394 AA.
AC      Q69091; 012544; 012833; P03171;
DT      21-JUL-1986 (Rel. 01, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Glycoprotein D precursor.
OS      GD OR US6.
SN      Herpes simplex virus (type 1 / strain 17).
OC      Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC      Alphaherpesvirinae; Simplexvirus.
OX      NCBI_TaxID=10299;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=85160822; PubMed=2984429;
RA      McGeoch D.J.; Dolan A.; Donald S.; Rixon F.J.;
RT      "Sequence determination and genetic content of the short unique region
RL      in the genome of herpes simplex virus type 1." ;
RL      J. Mol. Biol. 181:1-13(1985).
RN      [2]
RP      REVISIONS.
RA      McGeoch D.J.;
RL      Submitted (JAN-1989) to the EMBL/GenBank/DBJ databases.
CC      -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC      GB, GC, GG, GD, GI, AND GE.
CC      -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN D FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).

```

[illegible]

Query Match	10.5%	Score 99	DB 1	Length 394
Best Local Similarity	100.0%	Pred. No. 0.032		
Matches 19	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	159 QPELAPEDPEDSALLDPV	177
Dd	290 QPELAPEDPEDSALLDPV	308

RESULT 7	
VGDL_HSV1A	STANDARD;
ID_VGDL HSV1A	PRT; 394 AA.

DT	01-JUN-1994 (Rel. 29, Created)
DT	01-JUN-1994 (Rel. 29, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)

```

OS Herpes simplex virus (type 1 / strain Angelotti);
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10301;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90324869; PubMed=2165127;
RA Izumi K.M., Stevens J.G.;
RT "Molecular and biological characterization of a herpes simplex virus
RL type 1 (HSV-1) neuroinvasiveness gene.";
J. Exp. Med. 172:487-496(1990).
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN D FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54361; CA38245.1; -.
DR PIR; A47627; A47627.
DR InterPro; IPR002896; Herpes_glycop_D.
DR Pfam; PF01537; Herpes_glycop_D; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 394 GLYCOPROTEIN D.
FT DOMAIN 26 339 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 340 364 POTENTIAL.
FT DOMAIN 365 394 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 365 381 ARG/LYS-RICH (HIGHLY BASIC; PROBABLY
FT SERVES TO ANCHOR THE GLYCOPROTEIN IN

```

```
FT FT THE MEMBRANE).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 394 AA; 43303 MW; 47DE3BC79BB28950 CRC64;

Query Match 10.5%; Score 99; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLEDPV 177
Db 290 QPELAPEDPEDSALLEDPV 308

RESULT 8
VGLD_HSVLP STANDARD; PRT; 394 AA.
AC P57083; P03171;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glycoprotein D precursor.
GN GD OR US6.
OS Herpes simplex virus (type 1 / strain Patton).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_Taxid=10308;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=83016630; PubMed=6289440;
RA Watson R.J., Weis J.H., Salstrom J.S., Enquist L.W.;
RT "Herpes simplex virus type-1 glycoprotein D gene: nucleotide sequence
and expression in Escherichia coli.";
RL Science 218:381-384(1982).
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN D FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02138; CAA26060.1; -
DR EMBL; J02217; AAA45785.1; -
DR PIR; A94268; VGBED1.
DR PDB; 1JMA; 26-SEP-01.
DR InterPro; IPR002896; Herpes_glycop.D.
DR Pfam; PF01537; Herpes_glycop.D; 1.
KW Glycoprotein; Transmembrane; Signal; 3D-structure.
FT SIGNAL 1 25. POTENTIAL.
FT CHAIN 26 394 GLYCOPROTEIN D.
FT DOMAIN 26 339 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 340 364 POTENTIAL.
FT DOMAIN 365 394 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 365 381 ARG/LYS-RICH (HIGHLY BASIC; PROBABLY
SERVES TO ANCHOR THE GLYCOPROTEIN IN
THE MEMBRANE).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 394 AA; 43346 MW; 052AB5F53033D5E CRC64;

Query Match 10.5%; Score 99; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLEDPV 177
|||||
```

```
Db 290 QPELAPEDPEDSALLEDPV 308

RESULT 9
MEL_APICC STANDARD; PRT; 70 AA.
AC MEL_APICC
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Melittin precursor.
GN MELT.
OS Apis cerana cerana (Oriental honeybee),
OS Vespa magnifica (Hornet), and
OS Vespa velutina nigrit thorax (Hornet).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_Taxid=94128, 202807, 202809;
[1]
RN SEQUENCE FROM N.A.
RP SPECIES=A.c.cerana, V.magnifica, and V.v.nigrit thorax;
RC TISSUE=Venom gland;
RA Shi W.J., Zhang S.F., Zhang C.X., Cheng J.A.;
RT "Cloning and sequencing of cDNA coding for prepromelittin of Polistes
hebraeus, Vespa magnifica and Vespa maculifrons.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Main toxin of bee venom with strong hemolytic activity.
CC Integrates into cell membranes and has multiple effects, probably,
CC as a result of its interaction with negatively charged
CC phospholipids. It inhibits well known transport pumps such as the
CC Na(+)-K(+)-ATPase and the H(+)-K(+)-ATPase. Increases the
CC permeability of cell membranes to ions, particularly Na+ and
CC indirectly Ca(2+), because of the Na(+)-Ca(2+)-exchange (By
CC similarity).
CC -1- SUBUNIT: Monomer and homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- SIMILARITY: BELONGS TO THE MELLITIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF487907; AA012201.1; -
DR EMBL; AF487910; AA012204.1; -
DR EMBL; AF487908; AA012202.1; -
DR Pfam; PF01372; Melittin; 1.
DR ProDom; PD014636; Melittin; 1.
KW Cytolysis; Hemolysis; Toxin; Signal; Amidation.
FT SIGNAL 1 21. BY SIMILARITY.
FT PROPEP 22 43 REMOVED BY A DIPEPTIDYLPEPTIDASE (BY
SIMILARITY).
FT CHAIN 44 69 MELLITIN.
FT MOD_RES 69 69 AMIDATION (G-70 PROVIDE AMIDE GROUP) (BY
SIMILARITY).
SQ SEQUENCE 70 AA; 7543 MW; DA6B17C086C9560C CRC64;

Query Match 10.4%; Score 98.5; DB 1; Length 70;
Best Local Similarity 59.0%; Pred. No. 0.0048;
Matches 23; Conservative 4; Mismatches 7; Indels 5; Gaps 1;

QY 1 MKFLVVALVFMVVVYSIYAR-----LPDGITKAGEDA 34
Db 1 MKFLVVALVFMVVVYSIYAAPPEPAPEAAEAADAEA 39
|||||
```

RESULT 10
MEL_POLHE

```

ID MEL_POLHE STANDARD; PRT; 70 AA.
AC P59261;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Melittin precursor.
GN MELT.
OS Polistes hebraeus (Paper wasp).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Vespoidea;
OC Vespidae; Polistinae; Polistes.
OX NCBI_TaxId=202806;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Shi W.J., Zhang S.F., Zhang C.X., Cheng J.A.;
RT "Cloning and sequencing of cDNA coding for prepro-melittin of Polistes
RL hebraeus, Vespa magnifica and Vespa maculifrons."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Has strong hemolytic activity. Integrates into cell
CC membranes and has multiple effects, probably, as a result of its
CC interaction with negatively charged phospholipids. It inhibits
CC well known transport pumps such as the Na(+)-K(+)-ATPase and the
CC H(+)-K(+)-ATPase. Increases the permeability of cell membranes to
CC ions, particularly Na+ and indirectly Ca(2+), because of the
CC Na(+)-Ca(2+)-exchange (By similarity).
CC -1- SUBUNIT: Monomer and homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- SIMILARITY: BELONGS TO THE MELITTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF487909; AA012203.1; -.
DR Pfam; PF01372; Melittin; 1.
DR ProDom; PD014636; Melittin; 1.
DR Cytolysis; Hemolysis; Toxin; Signal; Amidation.
FT SIGNAL 1 21 BY SIMILARITY.
FT PROPEP 22 43 REMOVED BY A DIPEPTIDYLPEPTIDASE (BY
FT SIMILARITY).
FT CHAIN 44 69 MELITTIN.
FT MOD_RES 69 69 AMIDATION (G-70 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
SQ SEQUENCE 70 AA; 7513 MW; DA70167086C9560C CRC64;
Query Match 10.4%; Score 98.5; DB 1; Length 70;
Best Local Similarity 59.0%; Pred. No. 0.0048;
Matches 23; Conservative 4; Mismatches 7; Indels 5; Gaps 1;
QY 1 MKFLVNVALVEMVYISYIYAR-----LPDGITKAGEDA 34
DB 1 MKFLVNVALVEMVYISFIYAPPEPEAPEAEADADA 39

```

```

OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=85267679; PubMed=24107772;
RA Mount S.M., Rubin G.M.;
RT "Complete nucleotide sequence of the Drosophila transposable element
RL copia: homology between copia and retroviral proteins.";
RL Mol. Cell. Biol. 5:1630-1638(1985).
RN [2]
RP SEQUENCE FROM N.A., SEQUENCE OF 2-10, AND ALTERNATIVE SPLICING.
RX MEDLINE=85240569; PubMed=2409449;
RA Emori Y., Shiba T., Kanaya S., Imuye S., Yuki S., Saigo K.;
RT "The nucleotide sequence of copia and copia-related RNA in Drosophila
RT virus-like particles.";
RL Nature 315:773-776(1985).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=89183629; PubMed=2538806;
RA Miller K., Rosenbaum J., Zbrzezna V., Pogo A.O.;
RT "The nucleotide sequence of Drosophila melanogaster copia-specific
RT 2.1-kb mRNA.";
RL Nucleic Acids Res. 17:2134-2134(1989).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM SHORT), AND MUTAGENESIS OF ASP-292.
RC TISSUE=Larva;
RX MEDLINE=90151630; PubMed=1689241;
RA Yoshioke K., Honma H., Zushi M., Kondo S., Togashi S., Miyake T.,
RA Shiba T.;
RT "Virus-like particle formation of Drosophila copia through
RT autocatalytic processing.";
RL EMBO J. 9:535-541(1990).
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=long;
CC IsoId=P04146-1; Sequence=Displayed;
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A11.
CC -1- SIMILARITY: Contains 1 CCHC-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04456; CAA28054.2; -.
DR EMBL; X02599; CAA26444.1; -.
DR EMBL; X02599; CAA26445.1; -.
DR EMBL; X02600; CAA26446.1; -.
DR EMBL; X02600; CAA26447.1; -.
DR EMBL; X13719; CAA31997.1; -.
DR EMBL; X54147; CAA38086.1; -.
DR PIR; A03324; OFFFCP.
DR MEROPS; A11.001; -.
DR FlyBase; FBgn0013437; copia\GIP.
DR InterPro; IPR001969; Aspprotease_site.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00098; zf_CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS00141; ASP_PROTEASE; FALSE_NEG.
DR PROSITE; PS50158; ZF_CCHC; 1.
KW Transposable element; Hydrolase; Aspartyl protease; ATP-binding;
KW Polyprotein; Alternative splicing; Polymorphism; Zinc-finger.
FT CHAIN 1 270 COPIA VLP PROTEIN (POTENTIAL).
FT CHAIN 271 1409 COPIA PROTEASE (POTENTIAL).
FT ZN_FING 230 247 CCHC-TYPE.

```

```

FT ACT SITE 292 292 PROTEASE (BY SIMILARITY).
FT VARSPPLIC 392 1374 Missing (in isoform short).
FT VARIANT 1265 1288 /FTID=VSP_005226.
FT VARIANT 1265 1288 STGGVLFKMFDFNLICNTKRON -> VOOGIYSKLLIL
FT VARIANT 1289 1409 SFVGIOQRDTQ (IN VARIANT COPIA-RELATED).
FT MUTAGEN 292 292 MISSING (IN VARIANT COPIA-RELATED).
FT CONFLICT 191 191 D->A: LOSS OF ACTIVITY.
FT CONFLICT 300 300 S->N (IN REF. 2; CAA26447).
FT CONFLICT 866 866 I->V (IN REF. 2; CAA26447).
SQ SEQUENCE 1409 AA; 162817 MW; BE89440763AA7691 CRC64;

Query Match
Best Local Similarity 10.2%; Score 96.5; DB 1; Length 1409;
Matches 38; Conservative 27; Mismatches 48; Indels 47; Gaps 8;

QY 19 IYARLPDGTITKAGEDALRPWKS-----TAKHPWFOIED--NRCYIDNGKLPARGSLVGN 70
DB 1016 IYMRLPQGISCSNDNVCKNKAIYGLKQAAACMFVFEQALKECEFN-----SS 1065
QY 71 MSRFVDPKADYGVGENIVV--HADDFEFPVGESESLKVN-----VNLDMPIFET 119
DB 1066 VDRCIY--ILDKNINENIYVLLVYDDVIATGDMTRMNNFKRYLMEKFMWTDLNEIKHF 1123
QY 120 LALRLVLOQDVIVLRCVPELRVDYTSAYM-----WNMQ 153
DB 1124 IGIRIEMQEDKTYL-----SOSAYVKKILSKFRME 1153

RESULT 12
VGLD_HSV2 STANDARD; PRT; 393 AA.
ID P03172;
AC 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein D precursor.
GN GD OR US6.
OS Herpes simplex virus (type 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10310;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84131549; PubMed=6321120;
RA Lasky L.A., Dowbenko D.J.;
RT "DNA sequence analysis of the type-common glycoprotein-D genes of
herpes simplex virus types 1 and 2.";
RL DNA 3:23-29(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=333;
RX MEDLINE=84159516; PubMed=6323270;
RA Watson R.J.;
RT "DNA sequence of the Herpes simplex virus type 2 glycoprotein D
gene.";
RL Gene 26:307-312(1983).
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=BBKC;
RA Terhune S.S., Coleman K.T., Sekulovich R., Burke R.L., Spear P.G.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
CC 2: GH, GB, GC, GG, GD, GI, AND GE.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN D FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC -----
DR EMBL; U12180; AAB60552.1; -
DR EMBL; U12182; AAB60554.1; -
DR EMBL; U12183; AAB60555.1; -
DR EMBL; K02373; AAA45842.1; -
DR EMBL; K01408; AAA45841.1; -
DR EMBL; AF021342; AAB72102.1; -
DR PIR; A03731; VGBED2.
DR InterPro; IPR002896; Herpes_glycop_D.
DR Pfam; PF01537; Herpes_glycop_D; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 25
FT CHAIN 26 393 GLYCOPROTEIN D.
FT DOMAIN 26 339 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 340 363 POTENTIAL.
FT DOMAIN 364 393 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 249 249 T->I (IN STRAINS 333 AND BBKC).
FT VARIANT 337 337 A->S (IN STRAINS 333 AND BBKC).
FT VARIANT 367 369 RSV->AQM (IN STRAINS 333 AND BBKC).
SQ SEQUENCE 393 AA; 43147 MW; A8514E21857AEDF2 CRC64;

Query Match
Best Local Similarity 9.6%; Score 91; DB 1; Length 393;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 159 QPELAPEDPEDSALLLEDP 176
DB 290 QPELVEDPEDSALLLEDP 307

RESULT 13
REPI_ZYGRO STANDARD; PRT; 410 AA.
ID REPI_ZYGRO
AC P13778;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DE Trans-acting factor B (REPI).
GN P.
OS Zygosaccharomyces rouxii (Candida mogii).
OC Plasmid psr1.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Zygosaccharomyces.
OX NCBI_TaxID=4956;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85210904; PubMed=3889347;
RA Araki H., Jeauripackul A., Tatsuami H., Sakurai T., Ushio T.S.H.,
RA Muta T., Oshima Y.;
RT "Molecular and functional organization of yeast plasmid psr1.";
RL J. Mol. Biol. 182:191-203(1985).
CC -1- FUNCTION: PLASMID PARTITION REQUIRE REPI, REP2, AND A CIS-ACTING
CC DNA SEQUENCE (KNOWN AS STB). REP 1 MAY ACT BY INTERCALATING IN
CC THE YEAST NUCLEAR MATRIX AND BINDING STB EITHER DIRECTLY OR VIA
CC REP2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

Best Local Similarity 22.4%; Pred. No. 1.9; Mismatches 76; Indels 86; Gaps 12;
Matches 55; Conservative 29;

QY 2 KFLVNVN-----LVFMVYISYIARLPDGTAKGADALRPW-----KSTAKPWFQ 48
DB 43 KGMVNIAMENRDYPIALIFVVKYLMKKLPYGVIPVNLEPEPYVLTNTILKRLKEHKEFA 102
QY 49 IEDNRCYIDN-GKLPF-----RGSIVGNMSR-----FV-----F 76
DB 103 NKDKEDFAERLHKLIAPDVSIPESRKDEILGQKKERVVTKTINENFLDPVNARPLQRF 162
QY 77 DPKADYGVGENTLYVHADVEFVP-----GESLKW-----NVRNLDVMPLEFETLRLV 125
DB 163 FEKLHNGTLVENLEVLCKVEILVSSKMLGGSFKLQIMAAVRELWV----- 210
QY 126 LOGDVIMRCVPELRVDY-----TSSAYMMNMQYGMVGO---PELAPEDPEDSA 171
DB 211 --GEMVCMWITNE--TDYGFDEGCGDDDEGSSVEVQNSQSASPGQDQEAQRAPEAPETSS 266
QY 172 LLEDPV 177
DB 267 QLFDKI 272

RESULT 14
CDGT_BACS2 STANDARD; PRT; 703 AA.
ID CDGT_BACS2 P31746;
AC P31746;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cyclomaltoextrin glucanotransferase precursor (EC 2.4.1.19)
DE (Cyclodextrin-glycosyltransferase) (CGTase).
GN CGT.
OS Bacillus sp. (strain 1-1).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29334;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-52.
RA Schmid G., Englbrecht A., Schmid D.;
RT "Cloning and nucleotide sequence of a cyclodextrin glycosyltransferase gene from the alkalophilic Bacillus 1-1.";
RL (In) Huber O., Szeftli J. (eds.);
RT Proceedings of the fourth international symposium on cyclodextrins, pp. 71-76, Kluwer Academic Publishers, Dordrecht and Boston (1988).
RL -1- CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation of a 1,4-alpha-D-glucosidic bond.
CC -1- COFACTOR: BINDS TWO CALCIUM IONS.
CC -1- SUBUNIT: Monomer.
CC -1- MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE MALTOLOGOSACCHARIDE PRODUCED.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC PIR; S26399; ALBSX1.
DR HSSP; P31797; 1CYG.
DR InterPro; IPR006589; Alp_amy1_cat_sub.
DR InterPro; IPR006048; Alpha_amy1_C.
DR InterPro; IPR006047; Alpha_amy1_cat.
DR InterPro; IPR002044; CBD 4.
DR InterPro; IPR006046; Glyco_hydro_13.
DR InterPro; IPR002909; IPT_TIG.
DR Pfam; PF00128; alpha-amy1ase; 1.
DR Pfam; PF02806; alpha-amy1ase_C; 1.
DR Pfam; PF00686; CBM_20; 1.
DR Pfam; PF01833; TIG_1.
DR PRINTS; PR00110; ALPHAAMYLASE.
DR PRODOM; PD001568; CBD_4; 1.
DR SMART; SM00642; Aamy_1.
DR SMART; SM00632; Aamy_C; 1.

KW Transferase; Glycosyltransferase; Calcium; signal.
FT SIGNAL 1 29
FT CHAIN 30 703
FT DOMAIN 30 160
FT DOMAIN 161 224
FT DOMAIN 225 428
FT DOMAIN 429 516
FT DOMAIN 517 600
FT DOMAIN 601 703
FT DISULFID 68 75
FT ACT_SITE 251 251
FT ACT_SITE 279 279
FT ACT_SITE 350 350
SQ SEQUENCE 703 AA; 78663 MW; 4D973FB21D0D9B0A CRC64;

Query Match 8.4%; Score 80; DB 1; Length 703;
Best Local Similarity 20.0%; Pred. No. 4.5; Mismatches 54; Indels 46; Gaps 5;
Matches 32; Conservative 28;

QY 16 ISYIARLPDGTITKAGEDALRPWKTSTAKHPF----- 47
DB 105 VENNVALHPSGTSYHGWARDYKKT--NPYYGNFDDPDRLMSTAHNGIKVIMDFTPNH 162
QY 48 ---QIEDNRCYIDNGKLPARGSIVGNMSRFEVDPKADYGVGENTLYVHADVEFVPE-S 103
DB 163 SSPALETNPVYENGAIDYNGALIGN-----YSNDQNLFFHNGGTDFSSYEDS 211
QY 104 LKMNVRNLDVMPLEFETLRLVLOGDVIML-RVPELRVD 142
DB 212 IYRNLYDLADYLNNTVMDQYLKESIKFMDKIGDIGRVD 251

RESULT 15
YJHT_ECOLI STANDARD; PRT; 368 AA.
ID YJHT_ECOLI P39371;
AC P39371;
DT 01-FEB-1995 (Rel. 31, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein yjht precursor.
GN YJHT OR B4310.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L., Blatner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.8 through 100 minutes.";
RT Nucleic Acids Res. 23:2105-2119(1995).
CC -1- SIMILARITY: Contains 7 kelch repeats.
CC -1- SIMILARITY: STRONG, TO H.INFLUENZAE HI0148.
CC -1- SIMILARITY: TO H.INFLUENZAE HI0148.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC -----
DR EMBL; U14003; AAA97206.1; ALT INIT.
DR EMBL; AB000501; AAC77266.1; ALT_INIT.
DR EcoGene; EG12562; yjht.
DR InterPro; IPR006652; yjht.
DR Pfam; PF01344; Kelch; 2.
KW Hypothetical protein; Kelch repeat; Repeat; signal; Complete proteome.
FT SIGNAL 1 19
FT CHAIN 20 368
HYPOTHETICAL PROTEIN YJHT.

```
FT REPEAT 40 84 KELCH 1.
FT REPEAT 86 137 KELCH 2.
FT REPEAT 139 173 KELCH 3.
FT REPEAT 174 219 KELCH 4.
FT REPEAT 222 265 KELCH 5.
FT REPEAT 287 336 KELCH 6.
FT REPEAT 338 367 KELCH 7.
SQ SEQUENCE 368 AA; 39572 MW; 1194F392C51EA204 CRC64;

Query Match 8.3%; Score 78.5; DB 1; Length 368;
Best Local Similarity 18.8%; Pred. No. 3;
Matches 36; Conservative 29; Mismatches 69; Indels 57; Gaps 9;

QY 18 YIYARL---PDGITKAGEDA-----LRPWKSTAKHPWFOIEDNRCYIDNGKLFARGSI- 67
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 88 YVFGIGKNSSEGLTQVFNDVHKYNPKTNSWYKLSHAPMGAGHYTFVHNGKAYVTGGVN 147

QY 68 -----VGNMSRFVFDPRADYGGVGENLYVHADVDFVFPGESLKMN 107
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 148 QNIFNGYFEDLNEAGKDSIAIDKINAHYFDKKA-----EDYFFNKFLISFDP-STQQWS 200

QY 108 VRNLDVMPIFEITLRLVLOGDVIML---RCVPELR-----VDYTSSAYMMNMOMYGMVG 158
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 201 YAG--ESPWYGTAGAAVWVKGDKTWILINGEAKPGLRTDAVFELDFTGNMLKMN----- 251

QY 159 QPELAPEDPED 169
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 252 --KLAPVSSPD 260
```

Search completed: January 22, 2004, 11:39:41
Job time : 27.5802 secs

